EL COLEGIO DE MÉXICO
CENTRO DE ESTUDIOS ECONÓMICOS

MAESTRÍA EN ECONOMÍA

TRABAJO DE INVESTIGACIÓN PARA OBTENER EL GRADO DE MAESTRO EN ECONOMÍA

OPTIMAL INCOME TAXATION AND REDISTRIBUTION: AN EMPIRICAL EXERCISE FOR MEXICO

EMMANUEL SALVADOR CHÁVEZ JIMÉNEZ

PROMOCIÓN 2011-2013

ASESORES:

RAYMUNDO MIGUEL CAMPOS VÁZQUEZ

GERARDO ESQUIVEL HERNÁNDEZ

JUNIO 2013

ACKNOWLEDGEMENTS

First, I want to thank my advisor Raymundo Campos for his wise advice and his patience, his tireless support and commitment made possible the completion of this work. I also want to thank my advisor Gerardo Esquivel for his wise advice and his support.

Second, I want to thank all my classmates of my Master's program, without the academic and moral support that I received from every one of them I could never have completed the program.

Third, I want to thank my mother, my father and my siblings, who have always been there to help me no matter what. Their lifetime support has made possible all my achievements.

Last but not least, I want to thank you Georgina. You have always been there when I need a wise council; I will always cherish your trust and company.

ABSTRACT

This work derives optimal marginal income tax rates for individual Mexican taxpayers. I use the non-linear income tax formula derived by Mirrless (1971). The formula depends on two key components: 1. taxpayers' income distribution, 2. the elasticity of taxable income with respect to the marginal income tax rate. Tax return data on individual Mexican taxpayers is not available for public use. If it were available, it would provide Mexican taxpayers' declared income, which would serve to obtain the two key components of the optimal marginal income tax formula. Thus, in order to obtain a reasonable value of the elasticity of taxable income, I use estimations of the parameter made for other countries. To obtain the income distribution of Mexican taxpayers, I use income data from ENIGH (a household survey). First, I assume that Mexican taxpayers are distributed as are all incomes in ENIGH, then I keep only those individuals in ENIGH that belong to the formal sector and use their incomes to construct the income distribution of taxpayers. Using the information described above I calculate optimal marginal income tax rates for the entire income schedule. I find that the optimal marginal income tax schedule follows a U-shape pattern where rates should be higher for low and top earners. I use these tax rates to calculate the effects of redistribution on the incomes of poor earners and on income inequality under different government redistributive tastes. After calculating optimal income taxes for the whole income schedule, I calculate optimal marginal income tax rates for top earners using the formula derived by Saez (2001); to do that, I get an estimate of the Pareto parameter of the income distribution. For reasonable values of the Pareto parameter and the elasticity of taxable income I find that marginal income tax rates for Mexican top earners should be around 45 and 60 percent.

CONTENTS

1. INTRODUCTION

There are large positive effects of having a secured, healthy and educated population,[1] and there is evidence that government can provide those services in an effective way.[2] In order to finance those services, government inevitably has to resort to taxation.[3] This brings an interesting problem for the government: individuals' optimal choice between taxable and non-taxable activities could be modified by changes in the income tax rate. If a government wants to raise revenue by increasing tax rates, taxpayers' response could be so large that government could end up collecting a smaller amount of revenue after the increase in tax rates.

Literature has modeled an optimal income tax rate that depends on the distribution of income and on the elasticity of taxable income, a parameter that captures all the relevant responses of individuals induced by changes in the income tax rate. This theoretical framework was developed by Mirrless (1971).[4] Economists have turned to this framework, and have estimated the empirical values of the parameters on which the optimal income tax formula depends. Based on these empirical estimates, they are able to derive non-linear optimal marginal income tax rates for different income levels.

The purpose of this research is to approximate the values of these parameters for the case of Mexico and plug these approximations in the optimal income tax formula derived by Mirrless (1971) to obtain optimal marginal income tax rates for the whole Mexican income schedule. I use the word "approximate" because data needed to estimate these parameters is not available for

[1] To review the effects of education on economic growth see Krueger and Lindhal (2001). To review the effects of health on economic growth see Bloom, Canning and Sevilla (2004).

[2] Lindert (2004) argues that governments have become very efficient in administering social expenditure. He also argues that the costs of social transfers financed by tax collection are essentially equal to zero due to the fact that there is no evidence that tax collection and social expenditure deter long-run economic growth.
Piketty and Saez (2012) show that the largest part of government social transfers are expenditure in education, health and pensions.

[3] As governmental provision of public services increased over the twentieth century, also did income taxation. An interesting review of the increase of government social expenditure across countries is presented by Lindert (2004). An equally interesting review of the increase in income taxation across countries is presented by Atkinson and Piketty (2007).

[4] Since Mirrless' (1971) first approach to optimal income taxation, authors such as Atkinson (1990), Toumala (1990), Saez (2001), Salanié (2003) or Piketty, Saez and Stantcheva (2011) have built on it. I discuss this with more detail in Section 3.

public use in Mexico.[5] Consequently, I will resort to international estimates to approximate the parameters for Mexico.

It is a large misfortune that Mexico does not have the relevant data available for public use because this places the government in an unfavorable position: as any other government, Mexican government needs to collect taxes in order to provide public services for the population, but researchers do not have access to the appropriate data to estimate individual taxpayers' behavioral responses caused by changes in the tax rate. Without these estimates, optimal income tax rates (i.e. tax rates that maximize revenue given taxpayers' behavioral responses and income distribution) cannot be calculated. Under this situation, definition of income tax rates will be at best defective, and it may be harmful for the economy.

My motivation to do this research is inspired by the problem above mentioned, but instead of abandoning the endeavor of deriving optimal income tax rates due to the lack proper data, I will attempt to derive those rates by making use of the work done for other countries and of available data for Mexico. Lack of data is the reason why there is no previous research that tries to derive optimal marginal income tax rates for Mexico. Thus, I hope this work opens the discussion of obtaining optimal marginal income tax rates in this country by making a first attempt to calculate them. A side result of this work will be a document containing a detailed description of the difficulties that arise when trying to define optimal marginal income tax rates (MITRs) in Mexico. I hope that the resulting description will draw the attention on the importance of guaranteeing public access to data needed to calculate them.

In order to calculate the rates, I need information on two main factors: the income distribution of taxpayers and the elasticity of taxable income with respect to marginal income tax rates. I approximate the elasticity of taxable income (ETI) based on estimations made for other countries. I make the assumption that the ETI for Mexican taxpayers is similar to that of other countries.

I get the distribution of incomes from ENIGH, which is a household survey. First, I assume that Mexican taxpayers' incomes are distributed as are surveyed incomes in all kind of activities in ENIGH. I should remark that this is a strong assumption for Mexico because of the large size of

[5] In order to estimate the elasticity of taxable income, the literature makes use of income declared in tax returns to measure how changes in the income tax rates influence declared taxable income. Tax returns are also employed to analyze taxpayers' income distribution.

the informal sector in Mexican economy. ENIGH, being a household survey, captures how incomes in all kinds of activities are distributed, including those in the informal sector. However, in order to get optimal MITRs, one should be interested in how actual taxpayers' incomes are distributed, and individuals in the informal sector, which are included in ENIGH, do not pay income taxes. This is the reason why I perform sensitivity analysis using only those individuals in ENIGH that belong to the formal sector and assume that their incomes are distributed as those of actual income taxpayers.

With the information described above I calculate MITRs for Mexican taxpayers. I get the same U-shape pattern as other researchers do for other countries:[6] poor[7] and rich taxpayers should pay higher taxes and middle income taxpayers should pay lower taxes. I find that the tax schedule depends highly on the ETI and on government redistributive tastes. I get large tax rates, I argue that this could be the result of, among others, endogeneity issues in my calculations that cannot be solved due to the lack of proper available data.

The Mirrless model considers that revenue collected from taxation is redistributed in a lump-sum manner. Using the tax rates I derive with the Mirrless (1971) formula, I present calculations that illustrate the impact of redistribution on incomes of the poor and on inequality. My results show that income tax systems have large positive effects on income equality and on incomes of poor individuals.

After deriving optimal MITRs for the whole income schedule, I derive them for top incomes. In order to calculate rates for high earners I need information on how they are distributed. This is measured by the Pareto parameter of the top tail of the income distribution.[8] Here, income data from ENIGH presents another problem: top incomes are greatly undersampled in household surveys. I solve this problem by inflating high incomes in ENIGH taking the example of how they are undersampled in household surveys in the United States. I do this by comparing top earners in a household survey with top earners in tax return data in the US. I obtain an upper bound measure of the Pareto parameter.

[6] See Saez (2001) or Brewer, Saez and Shephard (2010).
[7] The model derived by Mirrless (1971) considers that collected revenue is redistributed in a lump-sum manner, so if poor taxpayers pay higher taxes, their well-being is not necessarily affected after redistribution takes place. I describe this with more detail in Section 5.
[8] I explain the Pareto parameter with detail in Sections 4.2 and 5.4.

Using the information just described, I calculate optimal MITRs for top earners, my results show that if the elasticity of taxable income in Mexico is similar to what has been obtained for other countries, for reasonable measures of the Pareto parameter, MITRs for high income taxpayers should be around 45 and 60 percent.

Diamond and Saez (2011) argue that if a theoretical model is to be used to construct policy recommendations, it has to fulfill three conditions, it must be: 1) empirically relevant, 2) robust to changes, 3) implementable. In general, the optimal income taxation model discussed in this work fulfills all three conditions, thus it can be used to form policy prescriptions. Nonetheless, for the case of Mexico, due to the lack of relevant data, the first condition is not satisfied. Thus, a number of the results I get cannot be used to construct policy recommendations; they should not be taken as definitive income tax rates for Mexican taxpayers.

However, in spite of the problems related to access of relevant data, I get two results that an optimal marginal income tax system in Mexico should have. First, the optimal marginal income tax schedule should be U-shaped. Second, marginal tax rates for Mexican top earners should be around 45 and 60 percent. I get two additional strong results: 1. the positive effects of an income tax system with redistribution —as the one presented in this work— on low earners are large. 2. I get an upper bound measure of the thinness of the income distribution at the top.

The work proceeds as follows: in Section 2 I mention a brief history of income taxes in the world and in Mexico. In Section 3 I discuss the theoretical framework that is behind optimal income taxation. Section 4 presents data and methodology I use to derive marginal income tax rates in Mexico. In section 5 I present the marginal income tax rates I calculate, first I give MITRs for all incomes, including sensitivity analysis and numerical exercises that illustrate the effects of redistribution, and then I present calculations of MITRs for top earners. Section 6 concludes.

2. A BRIEF HISTORY OF INCOME TAXATION

History of taxes goes a long way back in human history. According to Salanié (2003), taxes appeared coupled with civilization in Egypt and in Mesopotamia since year 3,500 BC. However, the appearance of income taxes is more recent. England and other European states established their first income tax to finance the Napoleonic Wars (e.g. Britain's Pitt's Act of 1799, which

gave birth to the country's first income tax); nevertheless, these income taxes were abolished when the war ended. With the exception of England,[9] income taxes would reappear by the end of the nineteenth century or the beginning of the twentieth century in most modern states (see Atkinson and Piketty (2007)). France instituted its income tax in 1914, the United States did so in 1913. Canada created its first income tax in 1920. Germany (Prussia) created its first modern income tax in 1891. The federal government in Australia levied an income tax for the first time 1914 (although Australian colonies collected an income tax before that). Additional countries enacted their modern income tax systems for the first time in the same period of time: New Zealand in 1892, the Netherlands in 1915, Ireland in 1922, and Switzerland in 1915.

Not only did the twentieth century witness the rebirth of income taxes, it also saw a dramatic increase in their rate and progressivity in most industrialized countries. To take an example, the average income tax rate in France for rose from cero by the beginning of the twentieth century, to close to ten by the mid 1970's. This rate was, however, different for wealthier individuals, with average income tax rate for top 0.01% of French earners in the mid 1970's being close to 50% (see Atkinson and Piketty (2007)). Piketty and Saez (2012) indicate that income taxes became possible because most governments became able to monitor income flows and earnings in a modern economy.

By the end of the twentieth century, income tax rates for top earners were reduced in most industrialized countries; the reduction was more significant in English speaking countries. The United States is an example of this trend, in the 1960's the marginal income tax rate for top American taxpayers was 90 percent, by the 1980's it was reduced to 35 percent. Piketty and Saez (2007) show that similar trends in top income tax rate reduction appeared in the 1980's in other countries. Although in most industrialized countries income tax systems are less progressive now than in the mid twentieth century, income tax progressivity is a feature that still remains in most income tax schedules in rich countries.

Sierra (1959) describes the origins of the income tax in Mexico; it goes back to 1921 when a government decree known as "Ley del Centenario" established an income tax on wages and profits for individuals and companies. Although progressive, rates were small compared to present rates (they ranged from 1% to 4%). Income tax was consolidated in 1925 with the

[9] England reintroduced the income tax in 1842. A progressive income tax was set up until 1909 when the "super tax" was crated (see Atkinson and Piketty (2007)).

publication of "Ley del Impuesto Sobre la Renta" (Income tax law), this law reinforced the legal setting of the income tax in Mexico. The new law redefined income brackets and tax rates, and maintained the progressive character of the previous law. In 1939, "Ley del Impuesto al Superprovecho" was established; it notably increased the tax rate (from 6% to 16%) for top incomes. In 1941, the income tax law has again reformed, among other things; it increased the tax rate for individual top earners to 20%, and maintained the progressivity of the income tax rate. In 1946, the tax rate for top self-employed earners was increased to 30%.

Since then, several reforms have been made to the income tax law, income brackets and rates have changed, taxable and non-taxable activities have been redefined, however progressivity in income tax rates remains.

I will now proceed to describe the theoretical framework behind the literature of optimum income taxation. Once this framework is reviewed, I will move to examine the empirical findings based on that theory.

3. OPTIMUM INCOME TAXATION

Since income taxes are a relatively modern invention, it is not surprising that an optimal income tax theory was only recently developed. The first effort to create income tax formulas that consider the disincentive effects on labor supply induced by income taxation was performed by Mirrless (1971).[10] Since the disincentive effects on labor supply are a central feature of the Mirrless optimal income tax formula, before presenting the formula, I should review the effects that income taxes induce on labor supply, in particular, the income and substitution effects provoked by the income tax. For that purpose, I will introduce the standard model of labor supply as presented by Salanié (2003):

A consumer has an utility function $U(C, L)$ that depends positively on consumption C and negatively on labor L. If the consumer faces an income tax t, its budget constraint is $C \leq (1 - t)(\omega L + R)$. Where ω is the wage rate and R is nonlabor income. An increase in the income tax has two effects. First, it reduces income which reduces demand for leisure, and thus increases labor supply, this is the income effect. However, there's a second effect: the increase in the tax

[10] Several authors have built on the optimal income tax literature since Mirrless' first approach, among them are Saez (2001), Salanié (2003) or Piketty, Saez and Stantcheva (2011).

rate decreases the relative price of labor, therefore reducing its supply; this is the substitution effect. From the Slutsky equation, we know that the total change in the labor supply is the sum of the substitution and the income effects.[11] The substitution effect is always negative (i.e. if taxes increase then labor supply will decrease), consequently, if there's a sufficiently large positive income effect associated with the increase in the tax rate, increasing income taxes could lead to a rise in labor supply, however as we shall further see, data suggests that this is not the case: an increase in the income tax rate will tend to decrease labor supply.[12]

Let us now discuss the model of optimal income taxation developed by Mirrless (1971) which, as several authors argue,[13] is the first model that considers the key efficiency-equity tradeoff of taxation: the government has to collect taxes to redistribute income among society, and social welfare is bigger when incomes are distributed more equally, but taxation and redistribution also induce disincentives to work, which could result in decreasing collected revenue. Optimal taxes take into account these disincentives.

I will summarize the key components of Mirrless' (1971) model of optimal income taxation as presented by Salanié (2003): Consumers have heterogeneous productivities ω. Person ω earns a before tax income $Y(\omega)$ and pays a tax $T(\omega)$. Government cannot observe agents' productivities, so it has to tax observable income, thus $Y(\omega) = \omega L(\omega)$ and $T(\omega) = \omega L(\omega)$ where $L(\omega)$ is labor supplied by individual ω. The government has to choose an income tax schedule to maximize:

$$W = \int_0^{\infty} \Psi(\Phi(\omega)) dF(\omega) \tag{1}$$

Where $\Phi(\omega) = U\left(\omega L(\omega) - T\big(\omega L(\omega)\big), L(\omega)\right)$ is the after tax utility function of individual ω, i.e. individuals' utility function is defined over after tax consumption and labor. Function $\Psi(\cdot)$ represents redistributive tastes of the government by weighting individuals' utilities.[14] The maximization problem is subject to government's budget constraint:

$$\int_0^{\infty} T\big(\omega L(\omega)\big) dF(\omega) \geq R \tag{2}$$

[11] It is important to recall two elasticity concepts from microeconomic theory: the compensated elasticity only accounts for the substitution effect, while the uncompensated elasticity accounts for both income and substitution effects. Consequently, the uncompensated elasticity is the sum of the compensated elasticity and the income effect.
[12] In Section 3.2 we will review that labor supply is not highly responsive to changes in the income tax.
[13] See Diamond and Saez (2011), Brewer, Saez and Shephard (2010), Salanié (2003), Saez (2001), Atkinson (1995) and Tuomala (1990).
[14] I discuss more on government redistributive tastes below and in Section 4.2.

From these assumptions, Mirrless (1971) derived a non-linear optimal marginal income tax formula, however, as noted by Salanié (2003) and Saez (2001), the process of solving for the optimal tax rate is quite difficult and not intuitive, the work of Mirrless (1971) has 141 numbered equations, so I will present the result for the optimal marginal income tax formula considering the quasi-linear utility case as presented by Salanié (2003) and discuss its implications.[15] As Salanié (2003) mentions, the quasi-linearity assumption implies that there are no income effects on labor supply when taxes increase, as we shall see, data shows that absence of income effects is not an unrealistic assumption.

The optimal marginal income tax formula is:[16]

$$\frac{T\prime(Y)}{1-T\prime(Y)} = \left(1 + \frac{1}{\epsilon_L(\omega_Y)}\right)\left(\frac{1-F(\omega_Y)}{\omega_Y f(\omega_Y)}\right)\left(1 - \frac{D(\omega_Y)}{D(0)}\right) \tag{3}$$

Where ω_Y represents the productivity that corresponds to income Y.

As Atkinson (1995) mentions, efficiency is captured by the elasticity of labor supply ε_L because it recovers behavioral responses, and equity is captured by the density of the productivities' distribution $f(\omega_Y)$ and government's redistributive preferences $D(\omega_Y)$. Observe that the optimal marginal income tax formula is non-linear, as Brewer, Saez and Shephard (2010) explain, this "means that METRs at a particular point of the earnings distribution can be set to any value without altering METRs at other points" (Brewer et al (2010), pp. 101).[17]

I will proceed to discuss the formula in a more detailed manner and explain the intuition behind each one of its components. The term in the first brackets in the right hand side captures the behavioral response induced by a change in the income tax, measured by the elasticity of labor supply ε_L,[18] it is straightforward to see that, other things equal, the larger the elasticity of labor supply, the smaller the marginal income tax rate. This is an intuitive result, if labor supply is highly responsive to changes in the tax rate, then government cannot increase tax rates to

[15] For further inspection on derivation of the optimal income tax formula, refer to Salanié (2003), Saez (2001) or to Mirrless's (1971) original paper.

[16] As discussed above, the optimal MITR formula is derived as a process where government maximizes (1) subject to (2). This maximization process takes into account taxpayers' utility functions. The first order condition is derived with Hamiltonian optimization. Mirrless (1971) makes several assumptions for deriving the optimal income tax formula, among others: government has perfect information about taxpayers utilities, the costs of administering tax revenue are negligible, intertemporal problems are ignored, and differences in tastes across individuals are not considered, they only differ in their skills.

[17] METR stands for marginal earnings tax rate.

[18] Rigorously speaking, the elasticity of labor supply with respect to the marginal income tax measures the percentage change in labor supply caused by a percentage change in the marginal income tax rate.

increase revenue because tax payers will reduce their working hours and total revenue may decrease.

The term in the second brackets in the RHS is a remarkable result because it indicates that the optimal marginal income tax rate depends on the distribution of productivities among individuals. $f(\omega_Y)$ denotes de probability distribution over ω_Y, thus, as Atkinson (1995) and Brewer et al. (2010) mention, the term $(1 - F(\omega_Y))/\omega_Y f(\omega_Y)$ indicates that the marginal income tax rate should be higher where the density of taxpayers is small relative to the number of taxpayers with earnings exceeding this amount. Intuitively: if marginal tax rates increase at earnings ω_Y, since they are *marginal* rates, they will increase for all earnings above ω_Y (which are measured by $1 - F(\omega_Y)$), but this increase in taxes also creates a disincentive effect for earnings around ω_Y (which are measured by $f(\omega_Y)$), taxpayers around ω_Y may reduce their labor due to the increase in income taxes. If the gain from increasing taxes for earners above ω_Y is larger than the disincentive effects for earners around ω_Y then taxes should increase.

The last factor that defines the optimal marginal tax rate is the redistributive tastes of the government which are captured in the term inside the third brackets of the RHD. Function $D(\omega_Y)$ assigns weights to individuals' marginal consumption according to their productivities; therefore it takes different values according to the redistributive priorities of the government. If government values some productivity's ω_Y marginal consumption highly, then, other things equal, MITR for that ω_Y should be smaller. The MITR should be higher if government has bigger tastes for redistribution.[19]

I will now inspect in a more detailed manner the influence of the distribution of productivities in the formula. I will also inspect the optimal MITR formula for high incomes, this formula gives an optimal marginal income tax rate for all individual taxpayers with earnings above some high earnings level $\bar{\omega}_Y$. Then, I will make a larger review on the elasticity of labor supply: I will discuss the theoretical framework behind that parameter (and, as we shall see, on the elasticity of taxable income), then I will perform a review of empirical estimations on the parameter.

[19] Government tastes for redistribution are explained with more detail in Section 4.2.

3.1 *Distribution of productivities and the special case of top incomes*

The distribution of productivities indicates that not only do income taxes depend on how people respond to changes in tax rates, but also on how productivities are distributed among taxpayers, this is a very interesting result for empirical applications because the distribution of productivities varies among societies. Even if elasticities of labor supply are similar in different countries, if the distributions of productivities among countries are different (which is most likely), the optimal marginal income tax schedule of these countries will differ. Since the distribution of taxpayers' productivities is not observable, authors such as Saez (2001), Gruber and Saez (2002) and Brewer et al. (2010), use taxpayers' income distribution to calculate optimal tax schedules, results show that tax rates are sensitive to changes in the income distribution.[20] The intuition of how optimal MITRs depend on the distribution of taxpayers' incomes is the same as what was discussed when I used the term "productivities": MITRs should be higher at points where the density of taxpayers' incomes is small relative to taxpayers with earnings exceeding that income level.

A special case of Mirrless (1971) formula is that of top earners. Saez (2001) presents the optimal marginal income tax rate for earners with incomes above z:[21]

$$T' = \frac{1-g}{1-g+\zeta^u+\zeta^c(a-1)} \tag{4}$$

Where g denotes the weight that government gives to top earners' marginal consumption, ζ^u and ζ^c are respectively the uncompensated and the compensated elasticities of earnings,[22] and a is the Pareto parameter (more on this parameter below). Saez (2001) shows that (4) can be obtained from the general Mirrless formula, thus formula (4) is consistent with the optimality of the MITR

[20] As noted by Saez (2001), changes in marginal tax rates may induce changes in the income distribution, thus, the ratio $1 - F(\omega_Y)/\omega_Y f(\omega_Y)$ in Mirrless formula is endogenous. We will discuss this problem with more detail in Section 5.1 where I present results for the optimal marginal income tax schedule for Mexico. Readers may see Saez (2001) for more information on this problem, or the considerations in Atkinson and Piketty (2007) about the influence of progressive taxation on the income distribution in different countries.

[21] A more recent top MITR formula is presented by Piketty, et al. (2011), their formula is more complete than Saez's (2001) formula because it allows top MITRs to be calculated not only taking into account earnings elasticities, but also tax avoidance elasticities and top earners' bargaining elasticities. Their model considers that a large part of top earners' income is not determined by their marginal product but by their bargaining power, so their income increases at the expense of lower earners. Top MITR rates under their model are larger than under Saez's (2001) model because tax rates depend positively on top earners' bargaining elasticities. Top income tax rates in Piketty's et al. (2011) model diminish top earners bargaining power. In this work I will only derive Saez's (2001) simpler formula since available data permits approximating the parameters in that formula and not those in Piketty's et al (2011) more complex formula.

[22] As we will later see, data indicates that income effects are negligible, thus one could set $\zeta^u = \zeta^c$.

formula presented here for all incomes. However, Saez (2001) also uses elasticities of earnings with respect to taxes to derive formula (4) in a simpler way.

Let me explain now what parameter a is: top tails of income distributions are well approximated by Pareto distributions,[23] this means that the probability that the income of some individual is greater than z is:

$$P[Income > z] = \left(\frac{z_m}{z}\right)^a \quad (5)$$

Where z_m is the minimum value of z. a is the Pareto parameter, it measures the "thinness"[24] of the top tail of the income distribution. a in formula (4) plays the role of $1 - F(\omega_Y)/\omega_Y f(\omega_Y)$ in formula (3), a tells us that the MITR for top earners depends on how incomes of top earners are distributed, as $1 - F(\omega_Y)/\omega_Y f(\omega_Y)$ tells us that the MITR for all earners depends on how all incomes are distributed.

Salanié (2003), Saez (2001) and Brewer et al. (2010) perform calculations with reasonable values for the elasticities and the Pareto parameter, and show their results. As expected, the optimal tax rate for top incomes is highly responsive to the thinness of the top tail distribution and to the values of the elasticities.

Let me now move to the other defining component of the optimal income tax formula: the behavioral response of taxpayers.

3.2. *The Elasticity of Labor Supply and the Elasticity of Taxable Income*

3.2.1. *Theoretical framework.*

A parameter that deserves special attention is the elasticity of labor supply because it captures taxpayers' behavioral responses induced by changes in the income tax rate. This parameter is a key determinant of how much a government can increase tax rates in order to increase revenue. Extensive analyses with United States data have been conducted to measure the elasticity of labor supply, these analyses evaluate the impact that different tax reforms have had on labor

[23] Authors such as Saez (2001), Goolsbee, Hall and Katz (1999), and the work compiled in Atkinson and Piketty (2007), use Pareto distributions to approximate the income distribution for top earners.
[24] I discuss the "thinness" of the top tail of the income distribution with more detail in Section 5.4.

supply. They find small values for these elasticities;[25] these findings imply that government can rely on increasing tax rates to increase revenue because taxpayers will not change in a meaningful manner their working time chosen. Thus, government can collect a higher amount of taxes for each worked hour, and since people will continue working (almost) the same time, the amount collected by the government will increase.

However, as Feldstein (1999) notes, the approach of examining the behavioral responses induced in taxpayers by changes in the tax rate only by looking at changes in labor supply is incomplete. Tax changes induce more behavioral responses apart from changes in labor supply, tax reforms can induce individuals to change their form of compensation, this means, individuals can keep working but change their compensation subject to tax, to tax-free compensation. Individuals could also alter consumption: they could change from tax-related consumption, to tax-free consumption (e.g. by increasing consumption subject to deductions). This is why Feldstein (1999) accurately notes that instead of looking at the elasticity of labor supply, we should be interested in the elasticity of taxable income to evaluate the effects of changes in income taxes.

Feldstein's (1999) argument is explained clearly by Goolsbee, Hall and Katz (1999): individuals maximize utility not only over consumption C and leisure L, they also choose nontaxable compensation E and nontaxable consumption D. If ω is the wage rate, then individuals taxable income is $\omega(1-L)-E-D$, i.e. tax related compensation minus tax-free compensation and tax-free consumption. Therefore, individuals maximize $U(C,L,E,D)$, subject to a budget constraint:

$$C(1-t)[\omega(1-L)-E-D] \tag{6}$$

Where C now stands for consumption subject to taxation, and t is the marginal income tax rate. Define $(1-\tau)=1/(1-t)$, and rewrite the budget constrain as:

$$C(1+\tau)=\omega(1-L)-E-D \tag{7}$$

[25] To see a broad exploration of labor supply response to tax changes, Pencavel (1986) offers a survey on estimates made by different authors that examine the responsiveness of working men to changes in income taxes in the United States and in the United Kingdom. MaCurdy (1992) reviews diverging estimates on the responsiveness of labor supply to tax changes; he concludes that raising income tax rates for high incomes causes small responses on men labor supply. Heckman (1993) makes a critique of the econometric procedures used by authors who estimate large responses of labor supply to income tax reforms.

The income tax changes the relative price of consumption subject to taxation and leisure but it does not change the relative prices of leisure, tax-free compensation and tax-free consumption. The approach of examining the elasticity of labor supply to measure the behavioral response of individuals when income tax rates change is therefore incomplete. When measuring taxpayers' responses due to changes in tax rates one should be interested in the substitution between taxable consumption[26] and tax-favored compensation and leisure. The parameter that measures such effect is the compensated elasticity of taxable income.[27] However, as Feldstein (1999) notes, changes in the income tax rate also induce income effects: if tax rates decrease, then disposable income increases, which induces the individual to choose more leisure and tax-free consumption, however Gruber and Saez (2002) find that income effects are highly insignificant; they state: "we can safely assume that compensated and uncompensated elasticities are identical and drop the income effect" (Gruber and Saez (2002), pp. 20). I will take this assumption throughout this work.

3.2.2. *Empirical Findings.*

In this subsection I discuss estimations of the elasticity of taxable income (ETI) in the United States.[28] I do this because it was in that country where literature on estimating the behavioral effects of taxation was constructed and where estimates were first carried out. In Section 4 I show estimates found for other countries which are mainly based on estimation strategies developed in the United States.

The first effort to estimate the elasticity of taxable income with respect to the marginal net of tax rate[29] was performed by Lindsey (1987). He uses separate cross-sectional data of high-income individuals' tax returns to analyze the behavioral responses to the tax reforms of years 1981-83 in the United States, which reduced marginal income tax rates. He assumes that data in different years corresponds to the same individuals. Lindsey concludes that the value of elasticity of taxable income is most likely in a range between 1.6 and 1.8.

[26] Or equivalently, taxable leisure, as Feldstein (1999) argues.
[27] By definition, the elasticity of taxable income is the percent decrease in reported taxable income when the income tax rate increases by 1 percent.
[28] An interesting review of the estimation of the elasticity of taxable income can also be found in Saez, Slemrod and Giertz (2012).
[29] The marginal net of tax rate is the net-of-tax income per marginal pretax peso.

A different method to estimate the ETI was carried out by Feldstein (1995), his method differs from Lindsey's (1987) in an important aspect, he uses panel data of over 4,000 individuals in the United States, this allows him to compare taxable incomes for the same individual before and after a change in tax rates. He examines the effect of the Tax Reform Act (TRA) of 1986, which reduced tax rates for the whole population, but reduction was sharper for high earners, who saw their marginal income tax rate decrease from 50 percent to 28 percent. He compares tax returns for 1985 and 1988. Feldstein estimates quite large elasticities of taxable income, ranging from 1.10 to 3.05 according to the taxpayer's income bracket.

Estimates performed by Lindsey (1987) and Feldstein (1995) indicate that individuals' taxable income is highly responsive to changes in the marginal tax rate. These results are discouraging for those trying to increase marginal tax rates in order to increase revenue: individuals' response to higher tax rates will be to substitute away a large part of their taxable compensation to tax-free or tax-favored compensation and consumption. However, Feldstein's results are criticized by Slemrod (1998) for having a small number of observations for high-income individuals (only 57), he states: "because of the wide variation among this group in financial situation and in income changes over time, generalizing over such small sample is problematic" (Slemrod (1998), pp. 775).

A considerable amount of literature examining the elasticity of taxable income surged following Feldstein's (1995) work. Auten and Carroll (1999) also examine the effects of the TRA of 1986 using panel data, but their dataset is larger, it consists of 14,102 tax returns and high-income tax payers are oversampled. They compare income tax reported in 1985 and 1989; their claim is that it is most likely that temporary responses to the 1986 tax rate change occurred before 1989, so the chosen time range eliminates temporary effects from ETI estimates. An essential difference of Auten and Carroll's (1999) analysis from Feldstein's (1995) is that they are able to include controls in their analysis, which allows them to identify non-tax factors that bias upwards the estimation of the behavioral response of taxpayers induced by changes in the tax rate. Without controlling for non-tax[30] factors and without weighting incomes to account for endogenous sample selection, they find an estimate for the elasticity of taxable income of 1.1, similar to that found by Lindsey (1987) and Feldstein (1995). The estimate of the elasticity of taxable income is lowered to 0.57 once they include controls and weights in the regressions.

[30] Non-tax factors included in their regressions are: age in 1985, age squared in 1985, entrepreneurship, number of children, and region of residence, among others.

Goolsbee (2000) takes a different approach for estimating the elasticity of taxable income, he estimates the effects of the tax increases of the Omnibus Budget Reconciliation Act (OBRA) of 1993, he uses data provided by Standard and Poor's to analyze the responses of highly-paid firm executives.[31] Data allows him to observe the same individual over time and to separate her income by source, which allows separating taxable compensation from tax-free or tax-favored compensation. A fundamental distinction of Goolsbee's (2000) study is that he is able to separate short-term responses from long-term responses. Goolsbee finds that large responses of taxable income to OBRA are "largely changes in the timing of the compensation, not in the form of the compensation" (Goolsbee (2000), pp. 366), he obtains an ETI estimate for highly-paid executives of above one for the short-term, but once he considers temporary effects, the elasticity estimate is less than 0.4. He also finds that wealthier executives have larger short-run elasticities, but long run elasticities are similar among executives.[32]

The 0.4 estimate is also found by Gruber and Saez (2002), although they analyze not only high-income taxpayers, but also taxpayers in the whole income range, their method has the advantage of examining taxpayers' responses for tax reforms of the 1980's and the 1990's, and that they use a large dataset containing nearly 100,000 observations, and as Auten and Carroll (1999), they are able to include controls. They define two types of income subject to taxation: taxable income (which is similar to the definition that other authors use), and broad income, they find that broad income elasticity is smaller than that of taxable income.

Giertz (2007) uses a similar approach to that of Gruber and Saez (2002) but with a larger dataset and he finds an ETI estimate of 0.3 for the period of 1971 to 2001. Auten, Carroll and Gee (2008) also use a similar approach to Gruber and Saez (2002) and Auten and Carroll (1999) to investigate the effects of the Economic Growth and Tax Reconciliation Act (EGTRRA) of 2001 and the Jobs and Growth Tax Relief Reconciliation Act (JGTRRA) of 2003, their ETI estimate is 0.39.

Another contribution to this literature are the historical examinations of taxable income elasticities performed by Gooslbee et al. (1999) and by Romer and Romer (2012). Golsbee et al. (1999) examine tax reforms in United States from the 1920's to the 1990's to get an ETI

[31] Goolsbee argues that the problem with this data is that executives may not be representative of all high earners, this may bias upwards the elasticity estimate.

[32] Goolsbee argues that this disparity is a result of wealthier executives receiving large compensations in form of stock options, which can be easily differed.

estimate. The quality of data is worse for decades before 1980, so they have to recur to different techniques to estimate elasticities such as statistical interpolations, they use this method to calculate the elasticity for the 1980's and find an estimate similar to Feldstein's (1995), they argue that this finding supports the validity of their method. The range found for the ETI estimate goes from negative values to close to 1.2, with most values ranging from nearly zero to 0.4. They claim that these values indicate that the elasticity of taxable income is not constant over time and its high values of the 1980's are an historical outlier. Romer and Rommer (2012) estimate the ETI for rich taxpayers in the United States in the interwar era. Their estimate is small; they find an ETI value of 0.2 for the whole period.

There range of estimates for the elasticity of taxable income found by different authors is wide, but this result may not be surprising. Saez, Slemrod and Giertz (2009) give two explanations for these differences, the first is that there may be issues in the estimation that some studies face and others do not, the second is that the ETI is not a structural parameter given by nature. Slemrod (1998) indicates some empirical issues when estimating the ETI. The first is tax rate endogeneity, there may be "exogenous variations in behavior that affect taxable income" (Slemrod (1998), pp. 781), when these factors are not considered the elasticity estimate may be biased. A second problem is due to changes in the definition of taxable income, tax reforms may change which types of compensation and consumption are subject to taxation, if this problem is not accounted for, the estimate could be biased in any direction.[33] A third habitual problem is of course available data, panel data is regularly not available for long periods of time, thus historical estimates of this parameter may not be accurate.

However, the second explanation may be more relevant for explaining why the elasticity of taxable income changes so much, as Slemrod states: "the elasticity of taxable income is a matter of government policy rather than a constant given by Nature" (Slemrod (1998), pp. 787). Law enforcement and base broadening influence taxpayers' behavioral responses. Consequently, government possesses instruments that can change taxpayers' choices between taxable and tax-free compensation and consumption, thus, the elasticity of taxable income "is not an immutable function of preferences" (Slemrod (1998), pp. 779). Change in government attitudes towards enforcement of tax law and changes in the definition taxable compensation induce behavioral

[33] Authors generally remark that they consider this factor in their studies.

changes in individuals which translate into changes in the elasticity of taxable income across time.

Taking into account all the difficulties associated with measuring the elasticity of taxable income, researchers agree upon an estimation of around 0.2 and 0.5,[34] those values are significantly smaller than the first estimates found by Lindsey (1987) and Feldstein (1995), nevertheless they indicate that taxpayers do respond to changes in the income tax rate, however, these elasticity values (depending on the shape of the income distribution), allow government to increase marginal tax rates in order to increase revenue if starting rates are low.[35]

4. DATA AND METHODOLOGY

4.1. *Data*

In order to calculate the non-linear optimum marginal income tax formula, as mentioned in Section 3, a researcher needs information on two factors: 1. the elasticity of taxable income, 2. the distribution of individual taxpayers' incomes.

4.1.1. *The elasticity of taxable income.*

In order to estimate the elasticity of taxable income, a researcher needs data on taxpayers' reported taxable income before and after changes in the income tax rate, this allows to observe how they respond to those changes, and thus, to estimate the corresponding elasticity. If information on declared taxable income is limited, i.e. there is only cross sectional data, one could estimate the elasticity of taxable income using a difference-in-differences method (as in Lindsey (1987) or Goolsbee et al. (1999)). If microdata is available, a researcher could use more robust methods to estimate the elasticity (as in Feldstein (1995), Auten and Carroll (1999) or Gruber and Saez (2002)). Neither cross-sectional data, nor panel data on income tax returns is available for public use in Mexico. I requested cross-sectional data that would allow estimating the ETI to Servicio de Administración Tributaria (SAT), the agency in charge of tax collection in Mexico. My request was denied so I filed an appeal at Instituto Federal de Acceso a la

[34] To see an illustrative description of different estimations of the elasticity of taxable income, consult Gruber and Saez (2002) pp. 5. A more recent review is presented in Saez et al. (2009).
[35] Bear in mind that these results are obtained for the case of the United States.

Información y Protección de Datos (IFAI), which is the institute that claims to guarantee access to public information in Mexico. IFAI's ruling favored SAT's information denial. As long as SAT does not provide the needed data, I will have to resort to a different procedure to approximate the ETI for Mexico.

There has been a recent surge in the estimation of the ETI. Not only in the United States have estimates been made, but also in a number of developed and developing countries. Table 1 presents estimates of this parameter made by various authors in different countries. The table shows estimates for 15 countries, I included the author's preferred point (or range) estimate in each study. Some interesting facts are shown.

First, the estimate of the ETI varies greatly not only among countries, but also among studies made for the same country; this is attributable to reasons mentioned in Section 3: the elasticity of taxable income is not a parameter given by nature; it is influenced by tax law and its enforcement, since governments and laws are remarkably different among countries, we should expect that the elasticities of taxable income be also different. In addition, the methods used to estimate the ETI vary across studies, some studies face some estimation issues that others do not, consequently, elasticity estimates for the same country may differ.

Table 1. Elasticity of taxable income estimates across countries

Country	Author	Earners subject to analysis	Time span	Data	Elasticity estimate					
					General	Self-employed	Wage earners	Highest earners	Male	Female
United States	Feldstein (1995)	All, except lowest earners	1985-1988	Panel data	1.1					
	Auteen, Caroll (1999)	All, except lowest earners	1985-1989	Panel data	0.57					
	Golsbee (2000)	Highest earners	1991-1995	Panel data				0.4-1.2		
	Gruber and Saez (2002)	All, except lowest earners	1979-1990	Panel data	0.4					
	Giertz (2007)	All earners	1971-2001	Panel data	0.3					
	Auten, Carroll, Gee (2008)	All earners	1999-2005	Panel data	0.39					
United Kingdom	Brewer, Saez, Shepard (2008)	Highest earners (richest 1%)	1978-2003	Cross-section				0.46		
Canada	Sillamaa, Veall (2001)	All, except lowest earners	1986-1989	Panel data	0.14-0.25	1.32		0.9-1.3		
	Saez, Veall (2005)	Highest earners (richest 1%)	1920-2000	Cross-section				0.2-0.4		
Germany	Gottfried, Schellhorn (2004)	All, except lowest and highly subsidized earners	1988-1990	Panel data	0.25-0.58	0.32	0.65-0.95	0.84-1.0		
Sweden	Hansson (2006)	All earners	1989-1992	Panel data	0.43-0.57					
	Blomquist, Selin (2010)	All earners	1981-1991	Panel data					0.24	0.9-1.4
New Zealand	Atkinson, Leigh (2008)	Highest earners (richest 1%)	1921-2005	Cross-section				0.41		
France	Piketty (1999)	Highest earners (richest 5%)	1985-1995	Cross-section				0.1-0.2		
Finland	Pirttila, Selin (2006)	All earners	1992-1995	Panel data	0.1-0.4					
Poland	Kopczuk (2012)	All earners, only business income	2002-2005	Cross-section / Panel data	0.95-1.29					
Pakistan	Kleven, Waseem (2011)	All earners	2006-2008	Panel data		0.1	0.07-0.12			
Denmark	Kleven, Schlutz (2012)	All, except highly subsidized earners	1985-2005	Panel data		0.1	0.05			
Spain	Díaz Mendoza (2004)	All, except lowest earners	1987-1994	Cross-section / Panel data	0.35					
Hungary	Bakos, Benczur, Benedek (2008)	All earners	2004-2005	Panel data	0.06			0.45		

This table was constructed based on estimations made by different authors of the elasticity of taxable income across countries. I include the preferred point (or range) estimate in each study.

Second, ETI estimates are similar in developed and developing countries. The average elasticity in the first group of countries is 0.42 and 0.41 in the second group.[36] Since tax evasion is greater in developing countries, a researcher may expect to find a higher ETI value in those countries, however, there might be a reason why that is not the case, as Kleven and Waseem (2011) indicate: "it is important to distinguish between the evasion *level* and the evasion *response* to marginal tax rates. A week enforcement system will be associated with large tax evasion, but not necessarily a large response of tax evasion to the marginal tax rate" (Kleven and Wassem(2011), pp. 19). This means: if taxpayers already evade large amounts of taxes, changes in the tax rate will not necessarily induce taxpayers to evade larger amounts.

Third, high-income taxpayers show higher ETI estimates compared with less affluent taxpayers in the same country. The average ETI estimate for high incomes is 0.57; this is 0.15 above the estimate for all earners (0.42). This indicates that top earners can more easily change their taxable income by moving to tax-favored consumption or tax-favored compensation.

4.1.2. *The distribution of taxpayers' incomes.*

It is a common practice across countries to guarantee public access to data that provides an estimate of the individuals' income distribution (as opposed to tax return data). Traditionally, the income distribution is estimated from income data obtained from household surveys, and Mexico is not an exception, however, the distribution of incomes can be obtained also from tax return data. The ideal income distribution source for deriving optimal MITRs would be incomes declared in tax returns, it would give a more precise estimation of how are taxpayers' incomes distributed. Again, the matter is that this information is not available for public use in Mexico, thus, in order to construct the distribution of Mexican taxpayers' incomes, I use a household survey. The survey I choose is ENIGH (Encuesta Nacional de Ingreso y Gasto de los Hogares), which is carried out yearly and is nationally representative. I will use data for year 2010, which contains information on income from 126,953 individuals. There are other surveys that contain information on individual income different from ENIGH; there is for example ENOE (Encuesta Nacional de Ocupación y Empleo), however, income information in ENOE is obtained from

[36] I must stress that there are few estimations for developing countries (only 4), of course this impedes to take this estimate as an average for all developing countries.

salary income, and income information in ENIGH is obtained from all sources of income. In view of the fact that not only salary income is subject to taxation, ENIGH can make a better approximation of the actual distribution of taxpayers' income.

By employing ENIGH for obtaining taxpayers' income distribution, I am implicitly assuming that taxpayers' incomes are distributed as individuals sampled in ENIGH. In Section 5.4 I present income distributions in the United States obtained both from a household survey and from tax return data, results show that for low and middle incomes, distributions obtained from both sources are very similar. Nevertheless, as opposed to the United States, a large part of Mexican labor force is employed in the informal sector, Levy (2008) indicates that in year 2006, 58% of Mexican work force was part of that sector,[37] and the informal sector does not pay income taxes, so assuming that taxpayer's incomes are distributed as incomes in all activities in ENIGH may be too demanding. Being a household survey, ENIGH samples individuals with in all types of activities, including individuals in the informal sector, since they are such a large part of the economy, the income distribution obtained from ENIGH might not be similar to the income distribution obtained from tax return data, which only takes into account taxpayers' incomes.

To face the issue of informality in Mexican economy, I perform sensitivity analysis using an income distribution taken from ENIGH which only considers those individuals that are part of the formal sector. I define individuals in the formal sector as those that have access to health services provided by the social security system,[38] in addition to these individuals, I also take employers as part of the formal sector. I present optimal MITRs obtained from both distributions, the one that takes individuals in all activities in ENIGH, and the one that takes only individuals that are part of the formal sector.

There is another problem related to income distributions obtained from household surveys: they undersample high earners, and this undersampling is quite large. This is the reason why Gruber and Saez (2000), Saez (2001), Brewer et al. (2010) and Díaz Mendoza (2004) use tax returns to

[37] Levy (2008) includes in the informal sector three types of individuals: 1. self-employed, 2. Individuals whose income is composed by commissions, 3. individuals that are employed by a formal firm but that, illegally, are not registered as such.

[38] Health institutions that are part of Mexican social security system are IMSS, ISSSTE, PEMEX and military health institutions.

obtain the distribution of taxpayers' incomes at the top of the distribution. Tax returns sample high-incomes better. Thus, using tax returns to obtain the income distribution increases the calculation's precision of the optimal MITR for high earners. This raises a problem when deriving optimal MITRs for top earners in Mexico: there is no public use data of tax returns that allows estimating accurately incomes at the top.

In order to face this problem, I compare the income distribution of the United States obtained with income data from household surveys with the one obtained with income data from tax returns.[39] I use the American Community Survey (ACS), it is the largest household survey carried out in the United States. I use year 2010, which contains data on over 2 million individuals. Income data in the ACS considers all sources of income, which is best for the purposes of this work. I compare the income distribution obtained from the ACS with the income distribution obtained from tax return statistics compiled by the Internal Revenue Service (IRS), and calculate the magnitude of the undersampling of high incomes in the ACS. Then, I apply this calculation to ENIGH; this allows making an approximation of the income distribution in Mexico that does not undersample top earners.[40] Here I implicitly make other assumption: household surveys' high-income undersampling in the United States is similar to household surveys' high-income undersampling in Mexico.

It is important to perform the comparing process described above if a researcher tries to calculate top marginal income tax rates in Mexico, as we saw in Section 3 and will see with more detail in the following subsection, the top MITR formula depends on how incomes are distributed at the top of the distribution. Table 2 compares average incomes for different ranges inside the top three percent earners in the United States measured by tax returns and by the ACS, it illustrates how greatly top incomes are undersampled in household surveys. Row P99.98-100 indicates the

[39] I use data for the United States because it is still not a common practice in the world to guarantee public access of individual tax return data. Most countries do not count with tax return data available for public use, others only give access to this data to national researchers, and others charge a fee if a researcher wants access to the data. It might be better to compare the income distributions of a developing country obtained from household surveys and from tax returns and take this result as a reference for Mexico, however, most developing countries have the same problem as Mexico: they don't grant public access to tax return data. Since international data on tax returns is highly limited, I use information on the United States following a common practice in economics: to take the US as a reference case.

[40] This method bears some resemblance to that used by Brewer et al. (2010). They merge the Family Resources Survey (a household survey), with the Survey of Personal Incomes (individual tax returns) to obtain the distribution of incomes. They use the FRS for incomes below £60,000, and de SPI for incomes above £60,000. Of course my method differs from theirs due to the fact that I do not use data on Mexican individual tax returns to obtain the distribution of top incomes.

average income of top 0.02% earners, according to tax return data, average income of top 0.02% earners on year 2010 was 16,286,145 dollars; ACS indicates that average income of top 0.02% earners was 752,287 dollars in the same year. For the United States, income of the highest earners is more than 21 times bigger if measured by tax return data instead of household surveys. This finding suggests that it is important to increase high-incomes in ENIGH when trying to derive top marginal income tax rates in Mexico in order to correct high income undersampling in household surveys.

Table 2
Average income of top-earners by source

Fractiles	Tax Returns	ACS
P97.00-98.00	225,986	156,517
P98.00-99.80	417,004	268,720
P99.80-99.90	1,229,950	459,091
P99.90-99.98	2,500,700	539,279
P99.98-100	16,286,145	752,287

This table shows average incomes for different ranges inside top 3 percent earners of the United States for year 2010. Average incomes for top earners are shown as obtained from the American Community survey and from tax returns compiled by the Internal Revenue Service.

4.2. *Methodology*

In order to calculate the optimal marginal tax schedule, I use the formula and notation presented in Brewer et al. (2010). The optimal non-linear marginal income tax formula is:

$$\frac{T'(z)}{1-T'(Z)} = \left(\frac{1}{e}\right)\left(\frac{1-H(z)}{zh(z)}\right)\left(1 - G(z)\right) \tag{8}$$

Where $T'(z)$ is the marginal income tax rate for taxpayer with income z. $H(z)$ is the cumulative distribution of taxpayers' incomes, i.e. the fraction of taxpayers with incomes less than z. $h(z)$ is the density of incomes. $G(z)$ denotes government's preferences of redistribution, i.e. how government values marginal consumption of taxpayers with incomes above z. The more government values consumption of individuals with income z, the less value $G(z)$ takes. If

government redistributive tastes, then $G(z)$ should be decreasing in z. Parameter e denotes the elasticity of taxable income.[41]

I approximate the value of the ETI in Mexico with the median of the elasticities estimated for countries shown in Table 1,[42] this value is 0.37;[43] I take this as the lower bound estimate for the ETI in Mexico. I add 0.3, and take that value as the medium estimation (0.67). My upper bound estimate is the median plus 0.6, i.e. 0.97. For the ETI of top earners I take the median of the ETIs for top earners shown in Table 1, and I repeat the same criteria: my lower bound ETI estimate for high incomes is 0.46, my medium estimate is 0.76, and my upper estimate is 1.06. I present calculations with the lower, medium and higher bound estimates of the ETI.

I obtain the distribution of incomes from ENIGH, as discussed in Subsection 4.1.2, I assume that the distribution of taxpayers is the same as the distribution of individuals in all types of activities in ENIGH, and then I perform sensitivity analysis using only individuals that are part of the formal sector of the economy.

Concerning government redistributive tastes, I make assumptions similar those made by Gruber and Saez (2000); I calculate the optimal marginal tax schedules with the following redistributive tastes:

- Rawlsian: a government with Rawlsian preferences is concerned on maximizing tax revenue in order to achieve the biggest income redistribution given taxpayers' behavioral responses, and it sets tax rates accordingly. In order to achieve maximum revenue, a Rawlsian government sets $G(z)$ equal to zero for all incomes: government does not value taxpayers' marginal consumption at any level, it is only concerned on income redistribution, and the higher revenue will allow the higher redistribution.

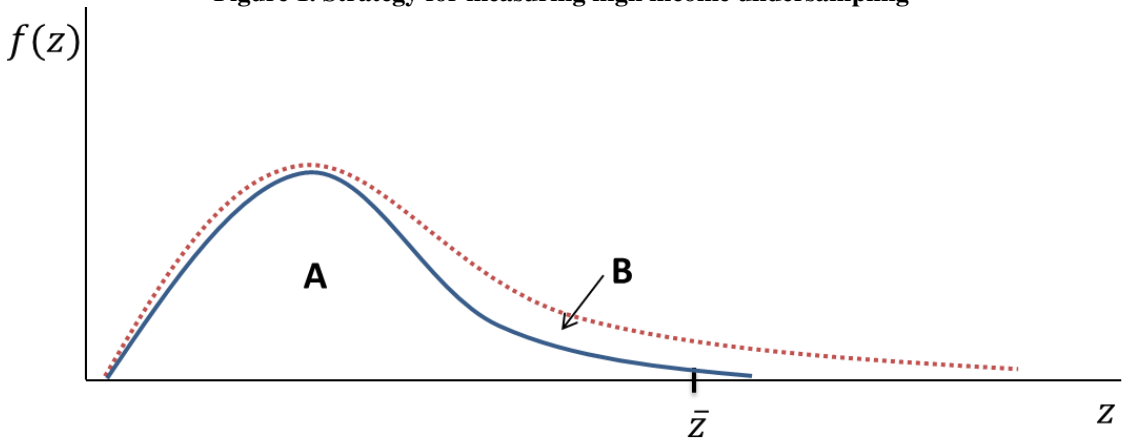
[41] Section 3 reviews the theoretical framework behind the optimal MITR formula and explains the intuition that lies behind each of its components. If the reader is interested in understanding the intuition behind the formula, I exhaustively recommend reviewing Section 3.

[42] This criterion for determining the elasticity of taxable income in Mexico makes an important assumption: it assumes that the ETI in Mexico is close to the median of ETIs in countries sampled in Table 1. There is no explanation that can validate this assumption. Mexico could be an outlier, the country could have an ETI above one or close to zero, and in that case, the marginal tax schedule obtained in this exercise would be bigger or lower than the optimal marginal tax schedule. However, given the lack of available data in Mexico, I found no better way to make an approximation of the ETI.

[43] An ETI of 0.37 means that a 1 percent increase in the marginal tax rate would decrease reported taxable income by 0.37 percent.

- Utilitarian: a government with utilitarian redistributive preferences values income redistribution, but it also values taxpayers' marginal consumption. I set preferences that give higher values to marginal consumption of lower incomes: government chooses a declining pattern for $G(z)$. There are two cases:
    - Progressive Liberal: The declining pattern of $G(z)$ is steep. The average value of $G(z)$ for small incomes is 1.5 times the average value of $G(z)$ for medium-small incomes. The average value of $G(z)$ for medium-small incomes is 3.4 times the average value of $G(z)$ for medium incomes. The value of $G(z)$ for high incomes is zero. Thus, as incomes get lower, government values more their marginal consumption.
    - Compassionate conservative: The government does not set different values of $G(z)$ for middle-low, middle and high incomes, but sets higher values of $G(z)$ for low incomes. This means that the government values equally marginal consumption of all taxpayers, except for the poor, their marginal consumption is relatively more valued.

Figure 1. Strategy for measuring high income undersampling



This figure presents the strategy I follow to estimate high income undersampling in ENIGH. Area A represents the income distribution obtained from household surveys. Areas A plus B represent the income distribution obtained from tax returns. Tax returns can provide a more accurate estimation of area B. I obtain area B by comparing the ACS with tax returns compiled by the IRS in the United States. Using ENIGH I obtain area A for Mexico, I add area B that I obtain comparing US income distributions to get area A plus B in Mexico.

When I calculate MITRs for high earners, as discussed Section 4.1.2, I merge the distribution of incomes of ENIGH with my measure of high-income undersampling. My strategy is illustrated in Figure 1, the solid line represents an income distribution obtained from household surveys (such

as the ACS), the dotted line represents an income distribution obtained from tax returns (as those compiled by the IRS), I compare IRS tax returns and the ACS to obtain area B. To get an income distribution for Mexico that does not undersample high incomes, I can get area A from ENIGH, and I add area B using the difference I obtain by comparing IRS tax returns and the ACS.

To obtain MITRs for high incomes, I use equation (4) and assume that there are no income effects to changes in taxation, thus: $\zeta^u = \zeta^c$.[44] Consequently, equation (4) can be rewritten as:

$$T' = \frac{1-g}{1-g+ae} \tag{9}$$

Where e is the elasticity of taxable income, g represents the value that government places to marginal consumption of high incomes, and a is the Pareto parameter of the income distribution. I obtain the Pareto parameter by calculating ratio $z_m/\bar{z}$ as done by Saez (2001), where z_m is average income above income level $\bar{z}$. Parameter $z_m/\bar{z}$ measures the "thinness" of the income distribution at the top. I first get ratio $z_m/\bar{z}$ using only income data form ENIGH, and then I compare income distributions of the US obtained both from the ACS and from tax return data. Then I apply this difference in incomes at the top to ENIGH, and I calculate ratio $z_m/\bar{z}$ again.

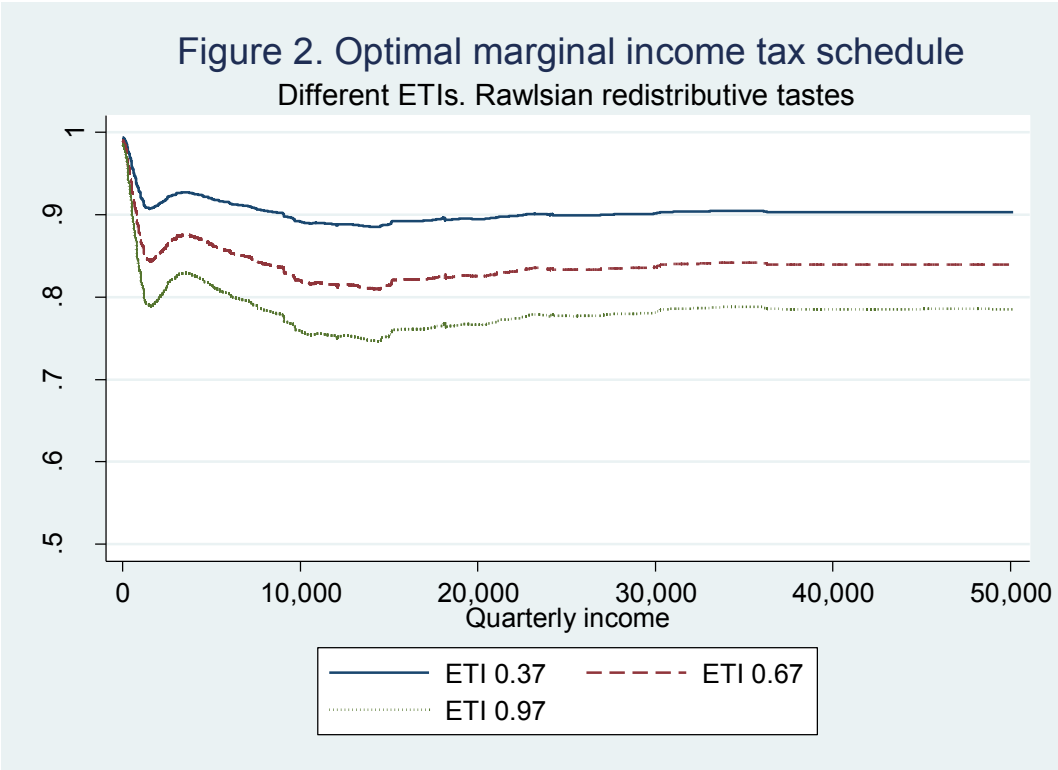
5. RESULTS

5.1. *Optimal Marginal Income Tax Schedules*

In this section I present results of my optimal MITR calculations obtained with formula (8) for the whole income schedule using data of individual earners in all types of activities in ENIGH. I start by showing the optimal marginal tax schedule when government has Rawlsian redistributive

[44] As reviewed in Section 3.2.2, empirical evidence finds that income effects are negligible, so it is a realistic assumption to set the compensated elasticity of taxable income equal to the uncompensated elasticity. If the reader is curious to see the influence of income effects in the top MITR formula, Table 1 in Saez (2001) presents calculations of the top MITR for the United States when income effects are considered, Saez shows that the optimal MITR for top earners is sensitive to changes in income effects. However, since empirical evidence finds negligible income effects, it is valid to set $\zeta^u = \zeta^c$ when deriving optimal top MITRs for Mexico.

tastes, i.e. it sets tax rates to maximize revenue and redistribution. Figure 2 shows these results. There are several facts that should be mentioned from results shown in Figure 2.[45]



This figure shows optimal marginal income tax schedules for Mexican taxpayers. Schedules are calculated using formula (8). The figure shows schedules for different values of the elasticity of taxable income. All estimations are made considering that government has Rawlsian redistributive tastes. Data for estimating the income distribution was obtained from ENIGH not distinguishing between individuals in the formal and in the informal sector. Values for the ETI were obtained from international estimates.

First, the optimal marginal tax schedule exhibits a U-shape pattern where marginal tax rates should be higher for low and high incomes, and lower for middle incomes. This U-shape pattern is also obtained for the United States by Saez (2001) and for the United Kingdom by Brewer et al. (2010). The pattern is driven by ratio $(1 - H(z))/zh(z)$, which summarizes the shape of the income distribution: "government should apply high marginal tax rates at levels where the density of tax payers, measured by $h(z)$, is low compared to the number of taxpayers with higher income, measured by $1 - H(z)$" (Brewer et al. (2010) pp. 105). This pattern results because ratio $(1 - H(z))/zh(z)$ measures two effects: if the MITR is increased at income z, then collection

[45] In Figures 2 through 6 I present optimal tax schedules with a cut-off level of $50,000 quarterly pesos. I do this because the optimal MITR "stabilizes" at that income level: for incomes above $50,000 quarterly pesos, the MIRT is the same.

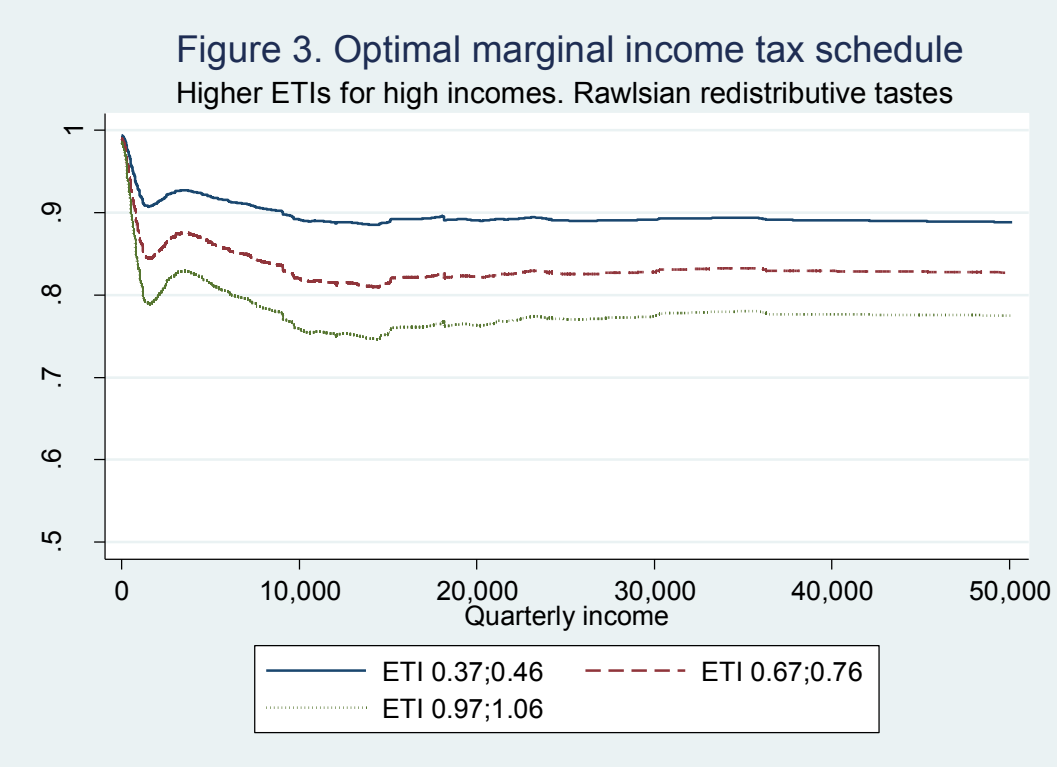
from all incomes above z will increase (because rates are *marginal*), but this also creates disincentives for incomes close to z (measured by $h(z)$), if the MITR increases at z, incomes close to z can be induced to hide away taxable income. So, if density at z is low compared to taxpayers with higher income, taxpayers subject to the disincentives of higher MITRs are comparatively smaller, this allows tax rates to be increased: the gains from increased revenue from all incomes above z is larger than the disincentives created for incomes around z.

Second, the ETI has a big influence on the optimal marginal tax rate. The average MITR is 92% when the ETI is 0.37, 87% when the ETI is 0.67, and 82% when its value is 0.97. This result is not surprising since we should expect that the more taxpayers' respond to changes in the income tax rate by shifting their subject-to-tax income and consumption to tax-free or tax-favored income and consumption, the lower the tax rates the government can set. Saez (2001) and Brewer et al. (2010) also find that ETIs affect the optimal MITR in a big way.

Third, the level of the optimal MITR can be shocking. Consider for example, MITR's for very low incomes, in all cases they are expropiatory: very poor individuals should pay 99% of their additional income in taxes. However, I should stress that the Mirrless (1971) model considers that the amount collected in taxes is redistributed to individuals in a lump-sum manner. Most likely, after collected revenue is redistributed to the population, poor individuals pay a negative average tax rate due to the fact that the redistributed income is larger than the amount they pay.[46] This is why a Rawlsian government cares only about the poorest members of the society: by setting $G(z) = 0$ for all z, the government maximizes tax revenue given the shape of the income distribution and the ETI. Other things equal, this will collect the highest amount of revenue which will in turn be redistributed to the population, this is for the benefit of the poorest individuals: their guaranteed income will be higher. As incomes grow, the amount of paid taxes becomes higher than the redistributed amount, thus average income tax rates become positive.[47]

[46] Even if collected revenue is not redistributed in a lump-sum manner, one could support the idea that a Rawlsian government makes the poorest individual better off: a higher amount of revenue allows government to provide more public services and social expenditure. The value of services provided by the government may be higher than the amount of taxes paid by poor individuals, resulting in poor individuals paying negative average income tax rates.

[47] Section 5.3 examines this with more detail.



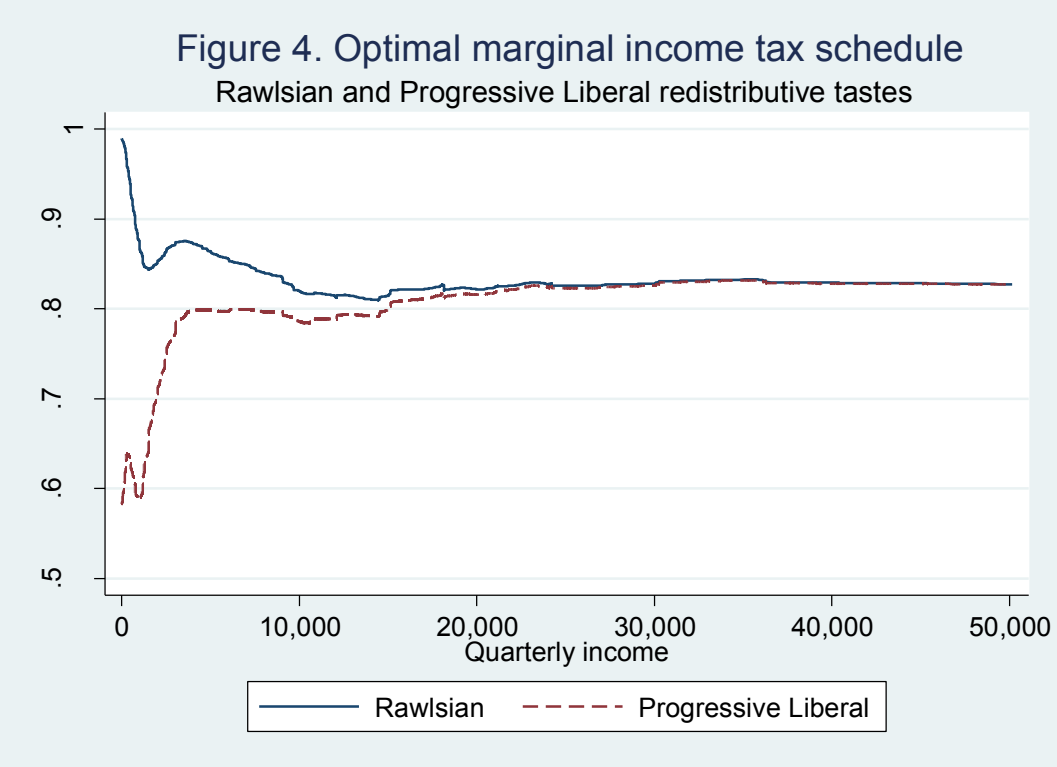
This figure shows optimal marginal income tax schedules for Mexican taxpayers. Schedules are calculated using formula (8). The figure shows schedules for different values of the elasticity of taxable income and considers than rich taxpayers hold higher elasticities. All estimations are made considering that government has Rawlsian redistributive tastes. Data for estimating the income distribution was obtained from ENIGH not distinguishing between individuals in the formal and in the informal sector. Values for the ETI were obtained from international estimates.

Marginal tax schedules shown in Figure 2 are derived assuming that all individuals bear the same ETI, however, as we have discussed, it is most likely that rich individuals have higher elasticities since they can use more tools to hide away taxable income. Figure 3 shows optimal marginal tax schedules when ETIs are higher for top earners. I calculate the schedules using the lower bound ETI estimate (0.37) for all incomes, and apply it to low and middle incomes, and the lower bound ETI estimate for high incomes (0.46). I do the same with the medium and higher bound ETI estimates.[48]

As expected, once I include higher elasticities for top earners, marginal income tax rates for richer taxpayers are lower. The U-shape pattern seen in Figure 2 is somewhat offset when I include higher ETIs for rich individuals, and the marginal tax schedule becomes and L-shape

[48] I assign the high income ETI to individuals with quarterly incomes above $27,100 Mexican pesos. The reason why I take this cut-off income level is because the marginal income tax rate "stabilizes" starting in this income: marginal income tax rates are almost equal for all quarterly incomes above $27,100.

pattern. As previously mentioned, it is most likely that the marginal tax schedule should follow a pattern that resembles to Figure 3 since we observe that top earners exhibit higher elasticities of taxable income across countries. As previously discussed, ETIs are higher for rich individuals because they count with more instruments to hide away or to change their compensation to tax-favored compensation (e.g. by increasing their share of compensation to stock options), acknowledging this circumstance, causes optimal MITRs for top earners to be lower.[49]



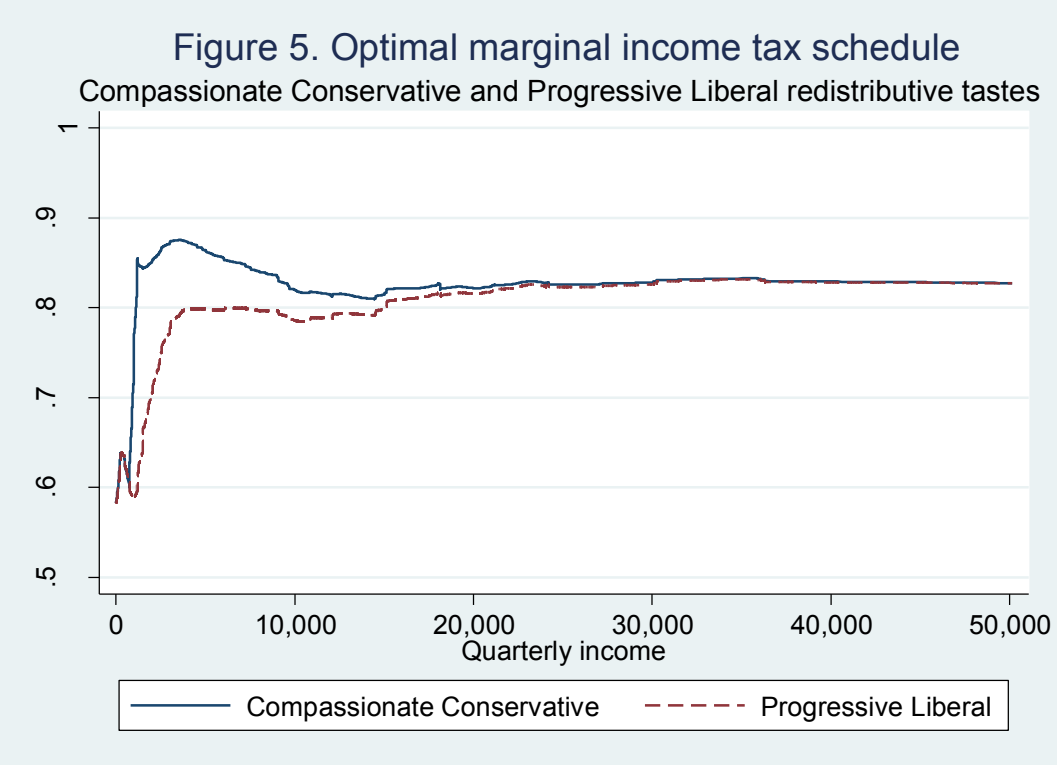
This figure shows optimal marginal income tax schedules for Mexican taxpayers. Schedules are calculated using formula (8). The figure shows schedules when government has different redistributive tastes. All estimations are made considering that the value of the elasticity of taxable income for taxpayers with low and medium incomes is 0.67, and 0.76 for high-income taxpayers. Data for estimating the income distribution was obtained from ENIGH not distinguishing between individuals in the formal and in the informal sector. Values for the ETI were obtained from international estimates.

An interesting exercise is to examine the large influence that government redistributive tastes have on the optimal marginal income tax formula. Figure 4 shows the marginal tax schedule when government has Rawlsian preferences compared with a Progressive Liberal government, in both schedules I use the mid-estimate of the ETI for low and middle incomes (0.67 for incomes

[49] Nevertheless, it must be kept in mind that the ETI is not a parameter given by nature; it is influenced by law definition and enforcement. If law does not allow high earners to recur to maneuverings to change their compensation when MITRs change, then ETIs for top earners should be smaller.

below $27,100), and the mid-estimate of the ETI for high incomes (0.76 for incomes above $27,100).

A progressive liberal government assigns very large weights to marginal consumption of low-income taxpayers; this drives optimal MITRs for poor taxpayers down. Since middle incomes also have a positive weight, their MITRs are also lower compared to a Rawlsian government. Only top marginal tax rates are equal because, as a Rawlsian government, a Progressive Liberal does not value marginal consumption of top earners.

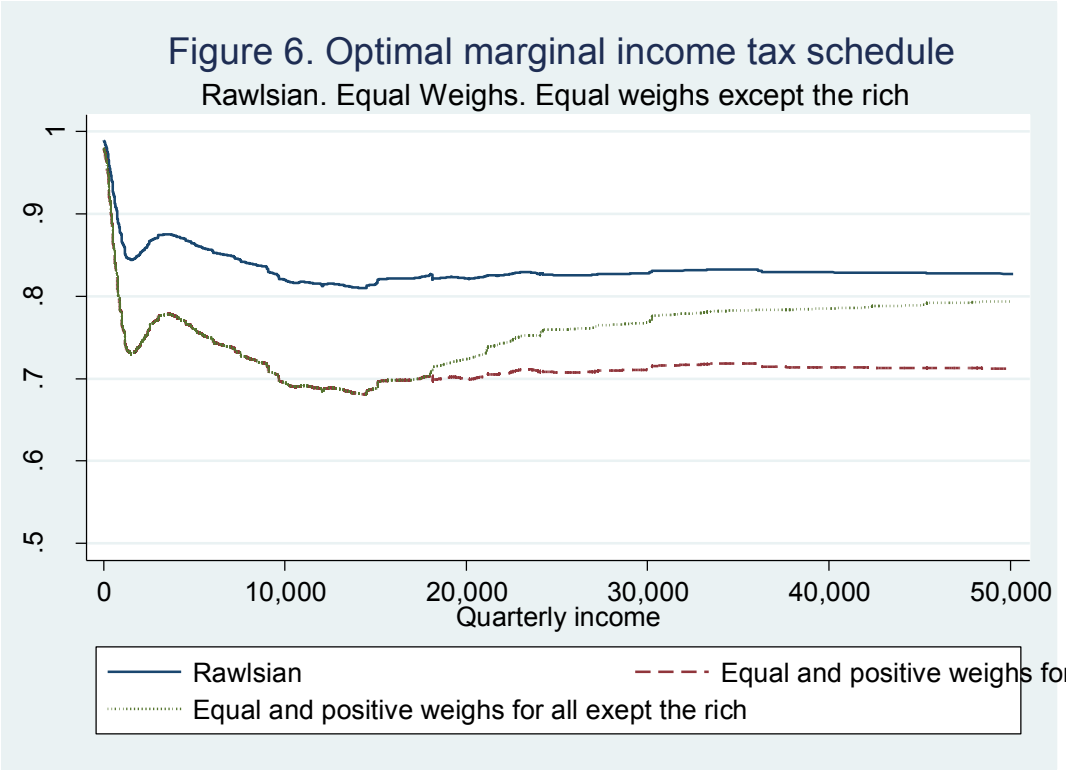


This figure shows optimal marginal income tax schedules for Mexican taxpayers. Schedules are calculated using formula (8). The figure shows schedules when government has different redistributive tastes. All estimations are made considering that the value of the elasticity of taxable income for taxpayers with low and medium incomes is 0.67, and 0.76 for high-income taxpayers. Data for estimating the income distribution was obtained from ENIGH not distinguishing between individuals in the formal and in the informal sector. Values for the ETI were obtained from international estimates.

It is important to note that poor individuals might be worse off with a Progressive Liberal government than with a Rawlsian government since revenue collected by the first is lower than by the second, therefore, the guaranteed income is lower in a Progressive Liberal government,

and after redistribution takes place, poor individuals might end having a lower after-tax income.[50]

Figure 5 compares optimal marginal tax schedules when government holds Progressive Liberal and Compassionate Conservative redistributive tastes. Interpretation of the results is similar to what we have previously discussed. We can see that MITRs for most of middle-low and middle incomes are higher in a Compassionate Conservative government, this is because the CC government assigns positive values to marginal consumption of very low incomes, and $G(z)$ for all other incomes is the same. In a Progressive Liberal government, $G(z)$ declines as income grows.



This figure shows optimal marginal income tax schedules for Mexican taxpayers. Schedules are calculated using formula (8). The figure shows schedules when government has different redistributive tastes. All estimations are made considering that the value of the elasticity of taxable income for taxpayers with low and medium incomes is 0.67, and 0.76 for high-income taxpayers. Data for estimating the income distribution was obtained from ENIGH not distinguishing between individuals in the formal and in the informal sector. Values for the ETI were obtained from international estimates.

A last and interesting comparison is between three different redistributive tastes: 1. a Rawlsian government, 2. a government that gives equal but positive weights to all taxpayers' marginal

[50] Section 5.3 clarifies this discussion with numerical examples.

consumption, 3. a government that gives equal and positive weights to all but the rich. Such a comparison is shown in Figure 6. A government that gives equal and positive weights to all incomes creates a schedule that mirrors a Rawlsian schedule; the only difference is that the former displays lower MITRs. It is more interesting to compare a government that values equally consumptions of all taxpayers, with a government that values marginal consumptions of all equally except for the rich (setting $G(z) = 0$ for high incomes). By comparing these two schedules, we can notice the great influence that government redistributive tastes have on the optimal marginal income tax formula. In the example shown in Figure 6, marginal income tax rates for very rich individuals can be ten percentage points higher if the government places lower values to their marginal consumption.

Table 3
Average marginal income tax rates by quarterly income level

		Poor	Middle-low	Middle	High	All incomes
		$0-$1,000	$1,001-$6,000	$6,001-$27,100	$27,101 or more	
		(1)	(2)	(3)	(4)	(5)
	Different Elasticity estimates*					
	0.37 for all incomes	96.6%	91.7%	89.8%	90.3%	92.5%
	0.67 for all incomes	94.1%	85.9%	82.9%	83.9%	87.4%
	0.97 for all incomes	91.7%	80.8%	77.1%	78.5%	82.9%
(1)	0.37 for low and medium, 0.46 for high incomes	96.6%	91.7%	89.7%	89.0%	92.5%
	0.67 for low and medium, 0.76 for high incomes	94.1%	85.9%	82.9%	82.8%	87.3%
	0.97 for low and medium, 1.06 for high incomes	91.7%	80.8%	77.0%	77.6%	82.8%
	Different Redistributive Tastes**					
	Rawlsian	94.1%	85.9%	82.9%	82.8%	87.3%
	Progressive Liberal	61.6%	70.8%	80.0%	82.7%	71.2%
(2)	Compassinate Conservative	62.8%	85.0%	82.9%	82.8%	77.8%
	Equal weighs of all incomes	89.1%	75.4%	70.8%	71.3%	78.0%
	Equal weighs of all incomes except top earners	89.1%	75.4%	71.4%	78.6%	78.4%

* Calculations obtained in row 1 consider a government with Rawlsian redistributive tastes.

** Calculations obtained in row 2 consider a value of 0.67 of the ETI for low and medium incomes, and 0.76 for high incomes

This table summarizes results obtained from optimal marginal income tax schedules for different elasticity values and different government redistributive tastes. Data for estimating the income distribution was obtained from ENIGH not distinguishing between individuals in the formal and in the informal sector. Values for the ETI were obtained from international estimates.

Table 3 presents a summary of marginal income tax rates for different values of the elasticity of taxable income and for different government redistributive tastes. I should retake the discussion on MITRs for poor individuals. The reason why MITRs are high for low earners when government does not value their marginal consumption is efficiency: the density of taxpayers is

high for low incomes, however, keep in mind that ratio $(1 - H(z))/zh(z)$ indicates that tax rates should be high at incomes where the density of taxpayers is low compared to taxpayers above that income, density of low income taxpayers may be high, but the relative size of the density of poor taxpayers is low compared to taxpayers with higher incomes (because *all* taxpayers have higher incomes). This means that it is efficient to raise taxes for low incomes because the disincentive effects created by raising taxes for the poor are little compared to the gains of increasing collection for all taxpayers that have higher incomes than the poor (since rates are *marginal*, increasing them for some income z, means increasing them for all incomes above z). In Section 5.3 I illustrate that it may be beneficial for their well-being if poor individuals pay higher income tax rates.

A remarkable result obtained from deriving optimal marginal income tax rates for individual taxpayers in Mexico is how high they can get. One could argue that this result is attributable to the shape of the income distribution in Mexico, however, I should point out that there are some issues with the calculations obtained in this Subsection:

1. The exercise is very simplistic, it considers that taxpayers only pay direct income taxes, nevertheless, if taxpayers also face other types of taxation such as consumption or payroll taxes (which surely is the case), then the optimal MITR must fall. As Salanié (2003) explains, other types of taxation also affect taxpayers' decisions of tax favored and non-tax favored compensation and consumption. Allowing for individuals to pay other types of taxation will change optimal MITR calculations.[51]
2. Note that changes in the marginal income tax rate can change the income distribution. Thus, there is an endogeneity problem with the optimal tax schedule derived in this work. The optimal MITR formula derived by Mirrless (1971) is not a function of taxpayers' income but of taxpayers' productivities. In the absence of income taxes, the income distribution should be equal to the distribution of productivities, and thus, the latter can be used to derive the optimal MITR schedule. If taxpayers do face income tax rates, a taxpayer's productivity can be approximated as a function of her income and her paid taxes. Thus, in order to calculate taxpayers' productivities, one needs data on taxes paid by

[51] Brewer et al. (2010) derive optimal MITRs for the United Kingdom considering a wide variety of taxes paid by UK taxpayers. I decided not to follow this approach due to the fact of limited availability of data on individual taxpayers in Mexico.

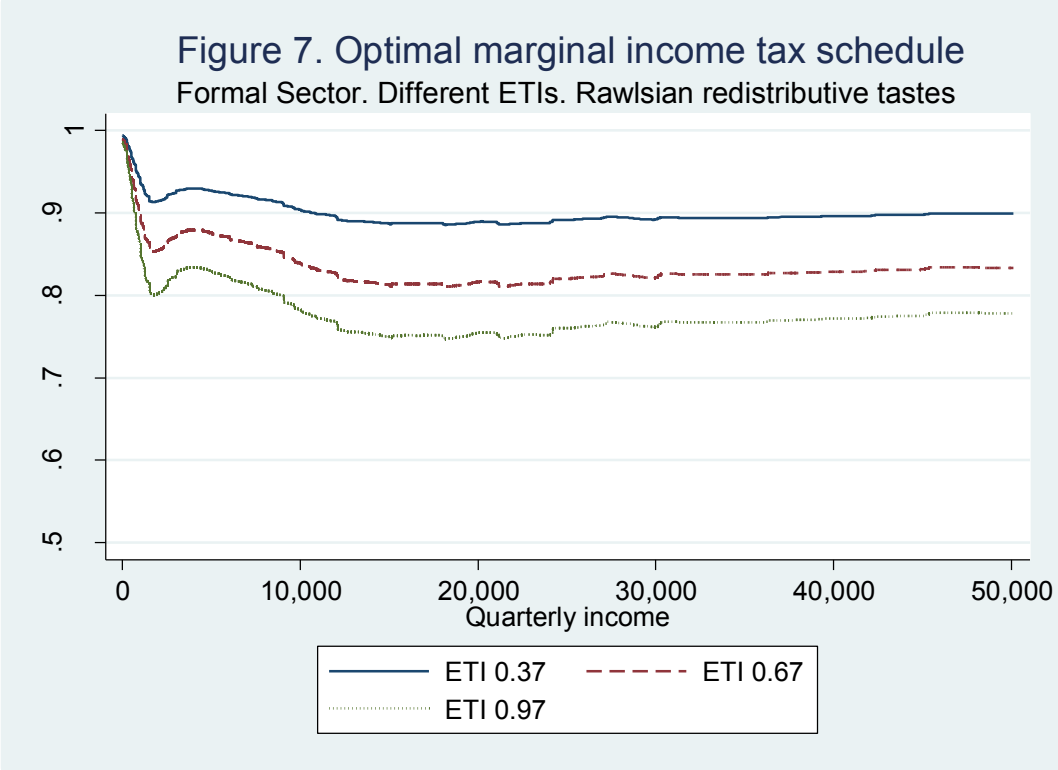
each individual. Since this kind of data is not available for public use in Mexico, I calculated the optimal MITR schedule by directly using the income distribution without adjusting it by taking in consideration taxes paid by taxpayers. As a consequence, my calculations bear an endogeneity problem.

3. An obvious problem is that the elasticities of taxable income employed when deriving optimal MITRs are not an actual result obtained from observing behavioral responses of Mexican tax payers to marginal income tax rates. Instead, due to the lack of available data, the ETIs employed here are obtained from estimations made for other countries' taxpayers. If actual Mexican ETI is significantly different from the values considered in this work, then MITRs are biased upwards or downwards.
4. Another problem in the schedules is the one discussed in section 4.1.2 about my assumption that Mexican taxpayers' incomes are distributed incomes in all types of activities surveyed in ENIGH. As argued before, this assumption is most likely valid for economies that do not have a large informal sector. Due to the fact that 58% of Mexican work force is employed in the informal sector, and the informal sector does not pay income taxes, assuming that incomes surveyed in ENIGH (where individuals that participate in informal activities are included) are distributed similarly to actual taxpayers' incomes might be invalid for the purposes of deriving optimal MITRs.
5. It should be kept in mind that the optimal tax schedule is a *marginal* income tax schedule. The tax schedule indicates how much of their additional income should taxpayers pay in taxes in the margin. An 80 percent marginal income tax rate does not mean that taxpayers should pay 80 percent of their total income in taxes; it means that taxpayers should pay 80 percent of taxes on their additional income.

5.2 *Sensitivity analysis*

In this subsection I present results of calculating formula (8) when I construct taxpayers' income distribution using only individuals sampled in ENIGH that are part of the formal sector. I only show results for the case when government bears Rawlsian redistributive tastes. Figure 7 shows these results. Results shown in Figure 7 are comparable to those shown in Figure 2. Note that optimal MITRs are very similar when I construct the income distribution from all incomes in

ENIGH and when I construct it using only incomes that are part of the formal economy: the schedule follows the same U-shape pattern and the level of tax rates is very similar.



This figure shows optimal marginal income tax schedules for Mexican taxpayers. Schedules are calculated using formula (8). The figure shows schedules for different values of the elasticity of taxable income. All estimations are made considering that government has Rawlsian redistributive tastes. Data for estimating the income distribution was obtained from ENIGH taking individuals that are part of the formal sector. Values for the ETI were obtained from international estimates.

I must stress that this resemblance does not mean that incomes in ENIGH are distributed as the actual distribution of taxpayers' incomes. ENIGH's sampling of individuals that are part of the formal sector may not approximate in a precise manner the distribution of individuals that actually pay income taxes. However, these results are very useful for evaluating the effects of redistribution that I show in the following Subsection.

The similarity that I get when estimating MITRs from income distributions that take and do not take into account individuals in the formal sector should serve as an argument favoring the importance that data on income tax returns be available for public use, since it would allow obtaining a more precise estimation of taxpayers' income distribution: there is nothing that

assures that incomes of individuals surveyed in ENIGH that are part of the formal sector are distributed in the same manner as incomes of actual taxpayers.

Note that optimal MITR schedules presented in this Subsection have the same problems as the schedules presented in the previous Subsection.

5.3. *Income Tax Redistribution*

As previously discussed, the Mirrless (1971) model of optimal income taxation considers that collected revenue is redistributed in a lump sum manner, so it is not necessarily harmful for poor earners if they pay a large share of taxes on their additional revenue, because after redistribution takes place, they almost surely will end up having a bigger income (as a consequence of guaranteed income being bigger than their income before redistribution). A consequence of the model is that well-being of poor earners depends on taxation of all earners, if taxes are high only for poor taxpayers, then collected revenue would be small and the redistributed amount might not be high enough to increase poor taxpayers' incomes. If taxes for middle and high earners are also high then collected revenue will rise and the redistributed guaranteed income will be big enough to increase poor taxpayers' incomes. Thus, high MITRs for poor earners are not prejudicial to their well-being only if collected revenue is high enough to provide a guaranteed income that is higher than their income before redistribution, and this will be the case only if middle and high earners are taxed high enough to provide that revenue (and thus, that guaranteed income).

In this Subsection, I present a numerical example that illustrates the effect that optimal income taxation under different government redistributive tastes has on incomes of the poorest earners and on the equality of the income distribution. I assume that collected revenue is redistributed in a lump sum manner to all earners and compare the effects that an income tax system would have under three different hypothetical economies: Economy 1 does not have an income tax system and the only income individuals get is their pre-tax income (this economy does not redistribute a guaranteed income). Economy 2 has an income tax system as the one presented in this work and a government that bears Rawlsian redistributive tastes. Economy 3 has an income tax system as the one presented in this work and a government that bears Progressive Liberal tastes.

First, I consider that all earners pay income taxes, so I use incomes from ENIGH in both the informal and the formal sector and calculate revenue under each tax system with tax rates presented in Section 5.1. Then I redistribute the collected amount in a lump-sum manner to all earners, each earner gets the same redistributed amount. I assume that the ETI for poor and middle earners is 0.67 and 0.76 for high earners. All earners pay income taxes and collected revenue is redistributed to all earners, thus there are not individuals that free-ride collected income since every earner pays a share of taxes. Table 4 shows some interesting results obtained from this exercise. Column 1 shows results for Economy 1. Column 2 presents results for Economy 2. Column 3 shows results for Economy 3.

Table 4
Effects of Redistribution on Poor Individuals. Only earners all types of activities pay income taxes.

	Economy 1	Economy 2	Economy 3
	No Income Taxation	Rawlsian	Progressive Liberal
Average income of the poor*	300.50	5747.43	5252.08
Share of the poor**	0.89%	17.08%	15.61%
Gini Coefficient	0.674	0.108	0.141

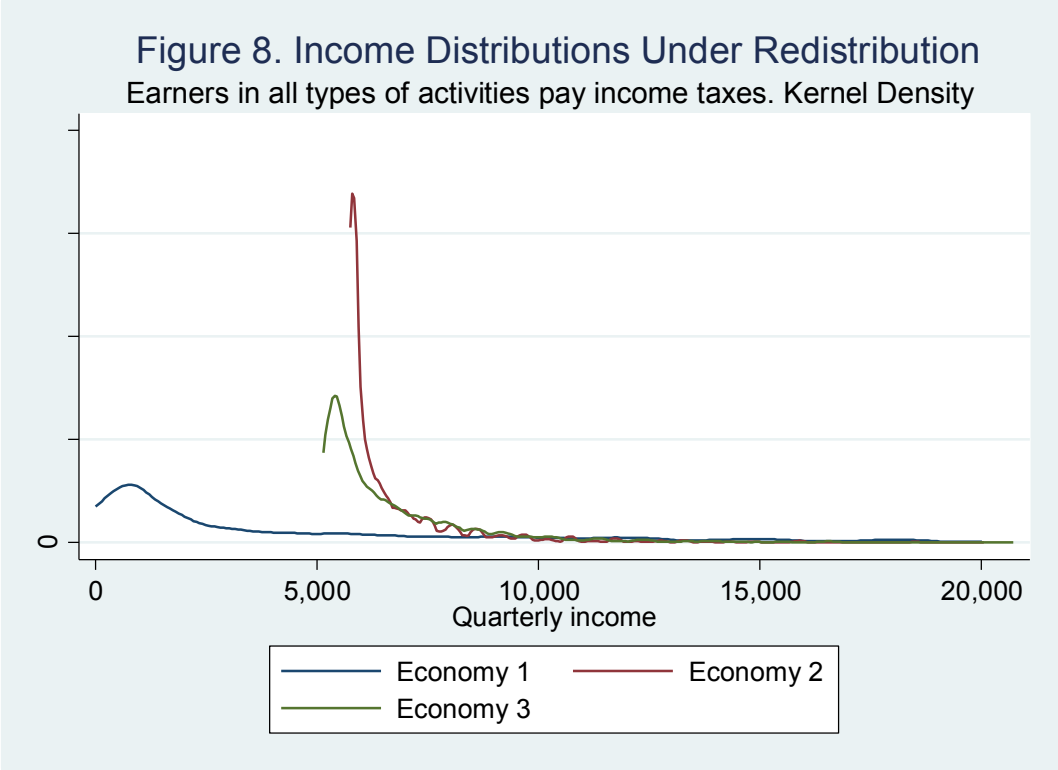
* Average quarterly income of the poorest 20% earners. (Mexican pesos)

** Share of 20% poorest earners in total income.

This table shows the effects of income redistribution on poor individuals. Collected revenue is redistributed by the government in a lump-sum manner. Government revenue is obtained from marginal income taxes shown calculated when I construct the income distribution without distinguishing between the formal and the informal sector. Every individual gets the same redistributed amount. The ETI for low and middle earners is 0.67 and 0.76 for high earners.

Results indicate the large influence that an income tax system that redistributes the collected amount in a lump-sum manner can have on incomes of poor taxpayers and on the income distribution. Average income of the poorest 20% earners is around 18 times larger in Economy 2 or Economy 3 compared to Economy 1, where the only source of income individuals get is their pre-tax income. The share of the poorest earners in total income is less than one percent in Economy 1, and it is more than 15 percent in economies 2 and 3. The Gini coefficient indicates

that the income distribution is considerably more equal in societies with an income tax system, reaching the highest equality when government has Rawlsian preferences.[52]

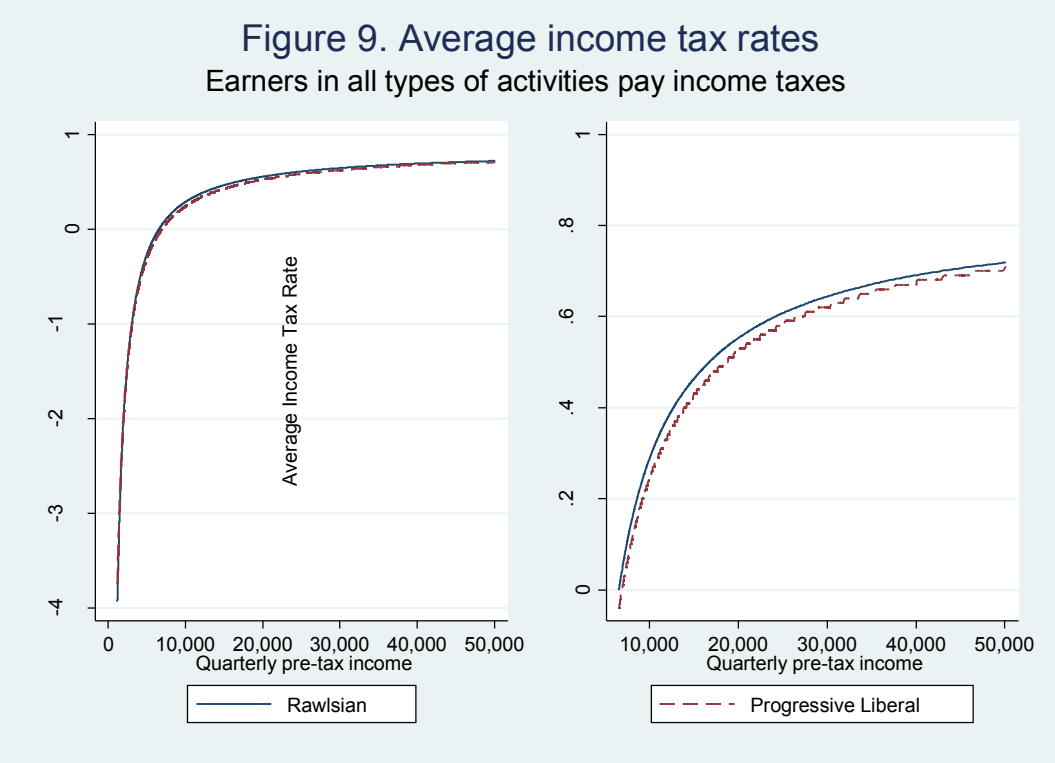


This figure shows different density estimates for the income distribution in societies where income taxation is present and societies where income taxation is not present. Distributions are constructed after redistribution takes place. Taxes are collected and redistributed under the assumption that all individuals pay income taxes.

Figure 8 shows Kernel density estimates for the three economies, densities are calculated after redistribution takes place. Densities shown in Figure 8 are consistent with what is shown in Table 4. The curve of the income distribution shifts upwards for low earnings once income tax systems are introduced.

Figure 9 shows average tax rates paid after redistribution taxes place. The fact that after-tax income for poor earners is bigger than pre-tax income results in poor taxpayers paying negative average tax rates. However, for middle and high earners, pre-tax income is bigger than after-tax income because redistributed guaranteed income is smaller than the amount of taxes they pay, as a consequence, middle and top taxpayers pay positive and increasing average tax rates.

[52] Hungerford (2011) performs an empirical analysis where he examines the relation between the Gini coefficient and income tax rates for high-earners in the United States; he finds that increasing income taxes reduces the Gini coefficient. His results point in the same direction as the results I get in my redistribution exercises.



This figure shows average tax rates paid after redistribution takes place considering that individuals in all types of activities pay income taxes. Average tax rates are obtained dividing the amount paid in taxes minus guaranteed redistributed income by taxpayers' taxable income.

Now I show results for the case when I consider that only individuals in the formal sector pay income taxes, so I use the MITRs I obtain when I plug in formula (8) the income distribution constructed only from earners that are part of the formal economy. Government redistributes collected revenue in a lump-sum manner not only to taxpayers but to all earners, thus in this example there are free riders that do not pay a portion of their income in taxes, but do receive a guaranteed income from the government when redistribution takes place.

Table 5 shows that, as in the case when all earners pay income taxes, an income tax system has a large impact on incomes of the poor and on income inequality when only earners in the formal sector pay income taxes. However, benefits for poor individuals are greater when all earners pay income taxes. When only earners in the formal sector pay income taxes, average income of the poorest earners is around 12 times bigger in economies 2 and 3 than in Economy 1 (compared to 18 times when all earners pay taxes). The share of the poor in total income also grows greatly when only earners in the formal economy pay income taxes, but not as much as when all earners pay. The Gini coefficient is also reduced.

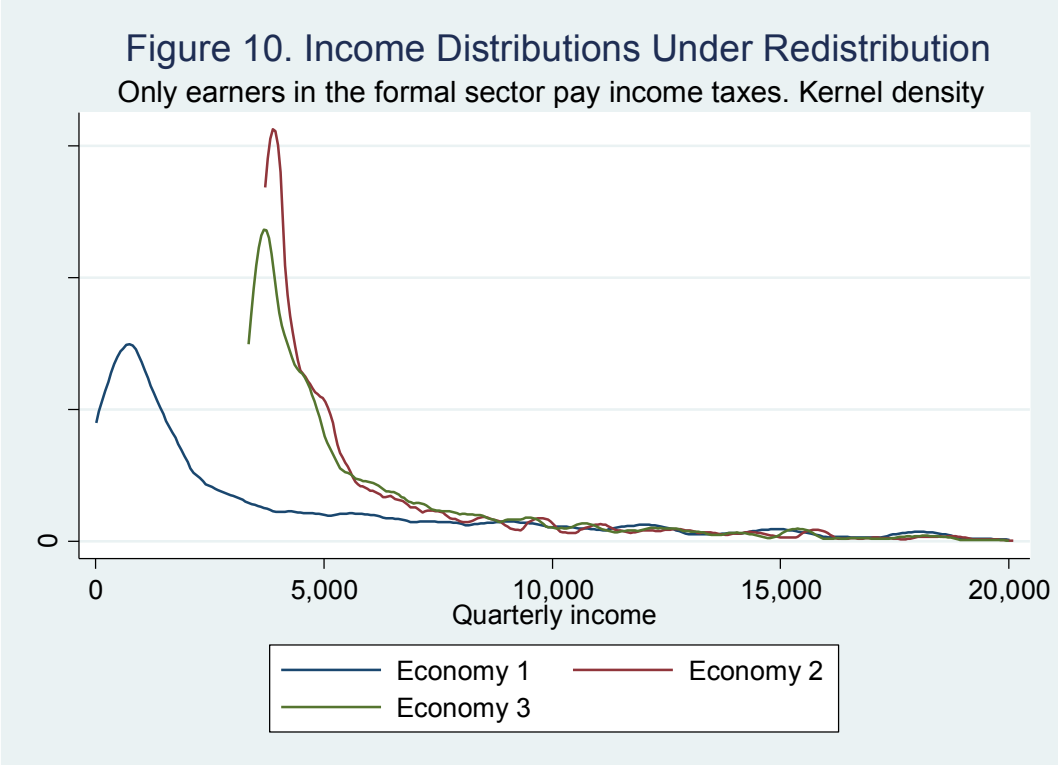
Table 5
Effects of Redistribution on Poor Individuals. Only earners in the formal sector pay income taxes.

	Economy 1	Economy 2	Economy 3
	No Income Taxation	Rawlsian	Progressive Liberal
Average income of the poor*	303.74	3848.12	3544.24
Share of the poor**	0.91%	11.47%	10.56%
Gini Coefficient	0.679	0.324	0.333

* Average quarterly income of the poorest 20% earners. (Mexican pesos)

** Share of 20% poorest earners in total income.

This table shows the effects of income redistribution on poor individuals. Collected revenue is redistributed by the government in a lump-sum manner. Government revenue is obtained from marginal income taxes calculated when I construct the income distribution only taking into account individuals in the formal sector. Every individual gets the same redistributed amount. The ETI for low and middle earners is 0.67 and 0.76 for high earners.



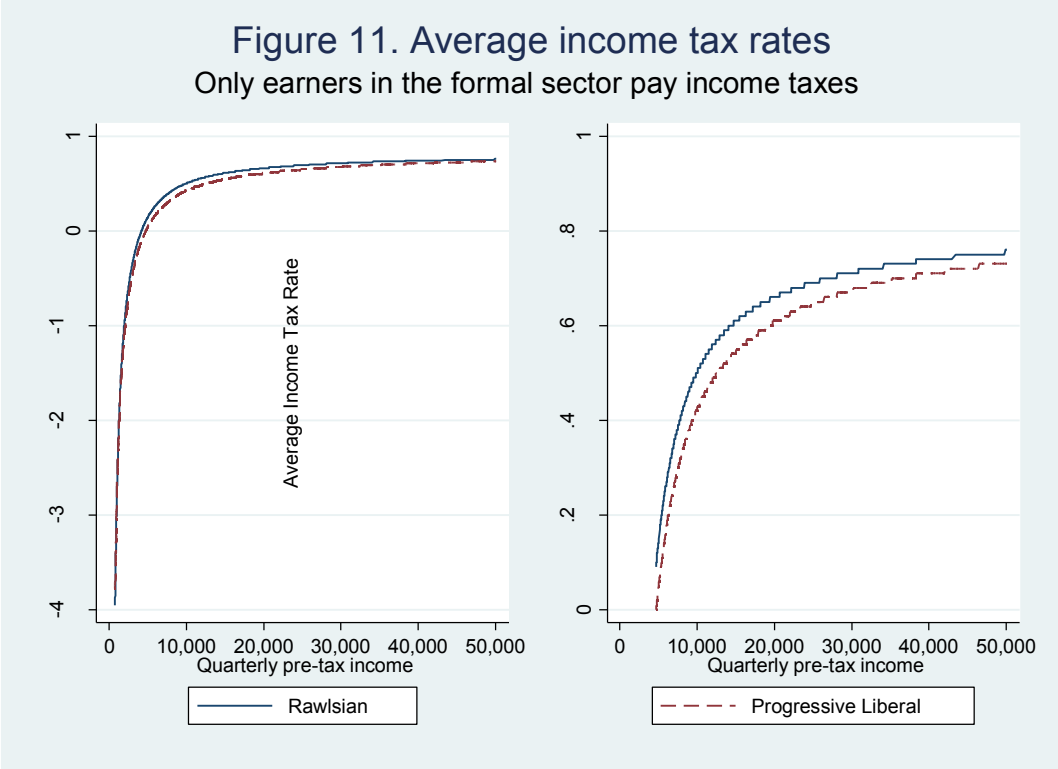
This figure shows different density estimates for the income distribution in societies where income taxation is present and societies where income taxation is not present. Distributions are constructed after redistribution takes place. Taxes are collected under the assumption that only individuals in the formal sector pay income taxes.

When only earners in the formal sector pay income taxes, in average, collected revenue per taxpayer is larger because the proportion of earners with high incomes is bigger in the formal sector, but the fact that there are less taxpayers when only earners in the formal sector pay

income taxes causes guaranteed income to be smaller, this causes that, in average, the benefits for poor earners are lower when only earners in the formal sector pay income taxes.

Figure 10 shows Kernel density estimates for the three economies, densities are calculated after redistribution takes place. In this case, as when all individuals pay income taxes, the curve of the income distribution shifts upwards for low earnings once income tax systems are introduced.

Figure 11 shows average income tax rates for taxpayers when only earners in the formal sector pay income taxes. Results are similar to the case when earners in all types of activities pay income taxes. Poor taxpayers pay negative average income tax rates, and as income grows, average tax rates become positive.



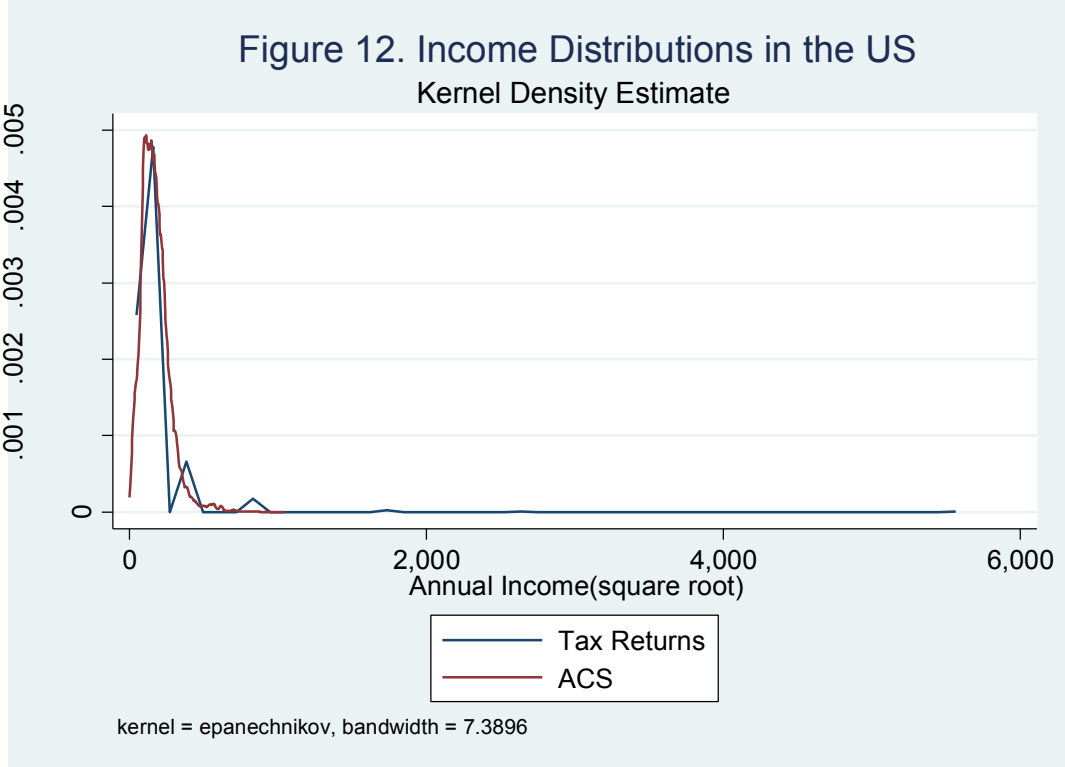
This figure shows average tax rates paid after redistribution takes place considering that only individuals in the formal sector pay income taxes. Average tax rates are obtained dividing the amount paid in taxes minus guaranteed redistributed income by taxpayers' taxable income.

Results shown in this Subsection support the idea that income tax systems that redistribute collected revenue have positive effects on the well-being of poor earners. They also support the argument that poor earners are better-off if government has Rawlsian redistributive tastes: poor

taxpayers pay larger MITRs, but, since collected revenue is also larger, the redistributed amount they get is bigger, and they end up having larger after-tax incomes.

5.4. *Optimal Top Marginal Income Tax Rates*

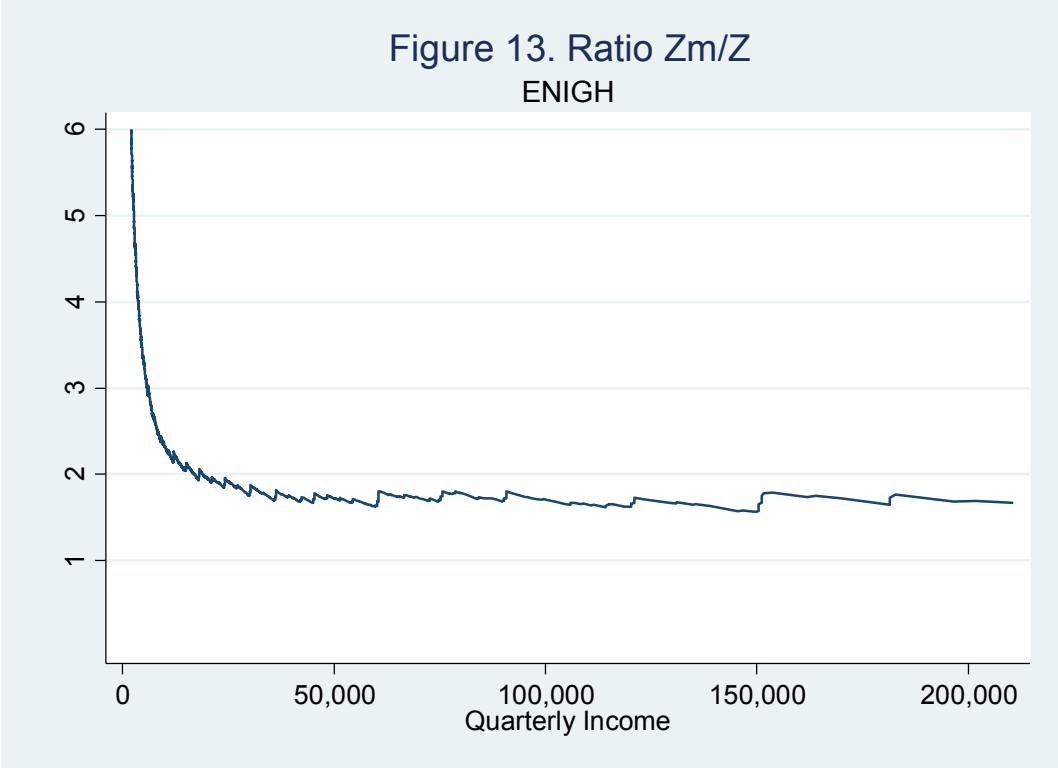
In this Subsection I present calculations of optimal marginal income tax rates for top earners, but before presenting them, I will show results of my estimations of the Pareto parameter. Figure 12 shows estimates of the income distribution of the United States obtained both from a household survey (ACS) and from tax return data. Figure 12 clearly shows how top incomes are undersampled by household surveys. Because of this undersampling, Pareto parameters obtained from household surveys are not reliable.



This figure shows the income distribution in the United States obtained from the American Community Survey, and from tax returns compiled by the Internal Revenue Service. Incomes are displayed in square root in order to highlight that concentration is higher in low incomes.

Figure 13 shows ratio $z_m/\bar{z}$ for the income distribution in Mexico obtained from ENIGH without inflating top incomes. Pareto distributions have the property that ratio $z_m/\bar{z}$ is the same for all $\bar{z}$ in the top tail of the distribution. Figure 13 shows that ratio $z_m/\bar{z}$ is more or less the

same for high $\bar{z}$'s. This confirms that the top tail of the income distribution in Mexico can be well approximated by a Pareto distribution (as happens in other countries). The Pareto parameter can be obtained from ratio $z_m/\bar{z}$, since $z_m/\bar{z} = a/(a-1)$. From a quarterly income of around $30,000 pesos and above, ratio $z_m/\bar{z}$ becomes stable at a value close to 1.7.[53] If $z_m/\bar{z} = 1.7$, then $a = 2.4$.

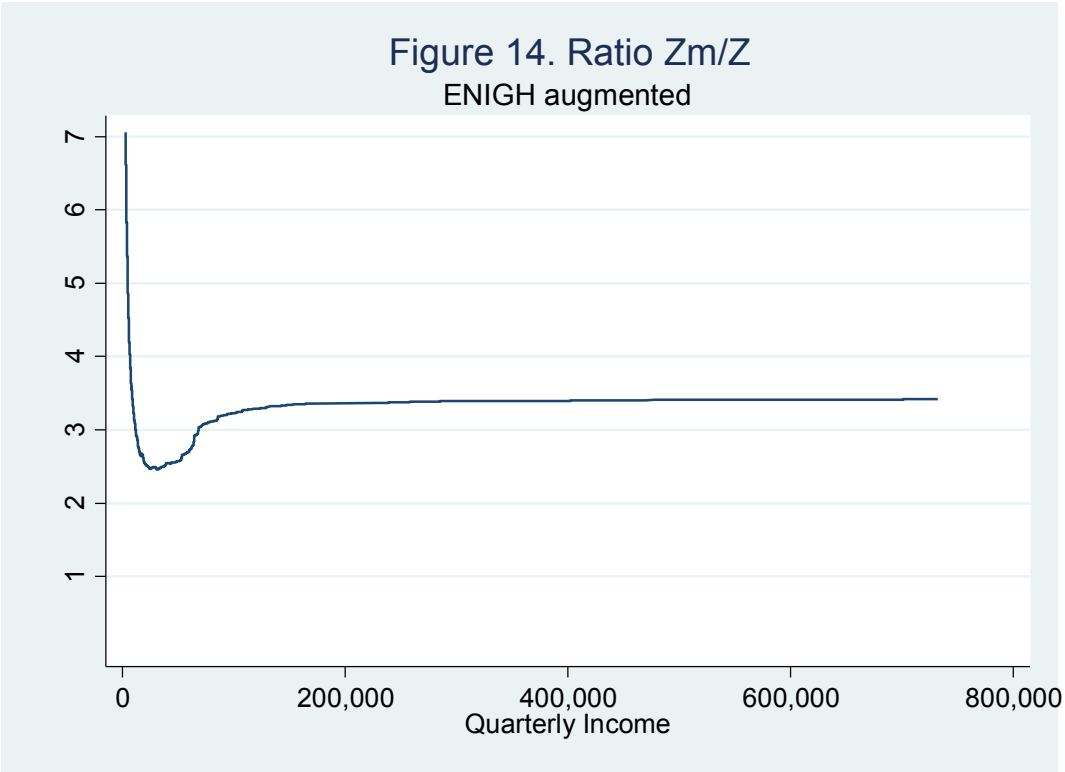


This figure shows ratio $z_m/\bar{z}$ for the income distribution in Mexico obtained from ENIGH when top incomes are not augmented to solve for high-income undersampling in household surveys.

Figure 14 shows ratio $z_m/\bar{z}$ for the income distribution in Mexico once I inflate high incomes in ENIGH taking the example of the differences between the ACS and tax return data for the United States. Ratio $z_m/\bar{z}$ stabilizes around quarterly income of $65,000 pesos, taking a value of about 3.5; this means that the Pareto parameter is around 1.4. The higher the Pareto parameter, the thinner is the top tail of the income distribution. It is straightforward to understand why the Pareto parameter is smaller once I augment top incomes in ENIGH: the top tail of the income

[53] Saez (2001) obtains ratio $z_m/\bar{z}$ for the United States, he gets the same stability pattern for high incomes. This empirical regularity was discovered by Vilfredo Pareto and is true for income distributions across countries.

distribution must be thinner if top earners are undersampled; consequently the Pareto parameter is smaller.[54]



This figure shows ratio $z_m/\bar{z}$ for the income distribution in Mexico obtained from ENIGH when top incomes are augmented to solve for high-income undersampling in household surveys.

Mexican data on tax returns would do a better job estimating the real value of the Pareto parameter, however, with calculations I perform in this work, I can set an upper bound to the parameter's estimate, note that ratio $z_m/\bar{z}$ calculated from ENIGH without inflating high incomes indicates that the Pareto parameter of the income distribution in Mexico cannot be bigger than 2.4, since this is the estimate obtained when top incomes are undersampled, the real value of the parameter must be smaller than 2.4, and most likely it is around 1.5 and 2. Atkinson, Piketty and Saez (2011) present historical international estimates of the Pareto parameter. For the decade of the 2000's, most values of the parameter range between 1.5 and 3. This means that the top tail of Mexican income distribution is fatter than that of many other countries.

[54] Ratio $z_m/\bar{z}$ obtained from ENIGH once top incomes are augmented follows the pattern of ratio $z_m/\bar{z}$ obtained by Saez (2001) for the United States more similarly than ratio $z_m/\bar{z}$ obtained from ENIGH without augmenting top incomes.

Table 6 presents results from calculating formula (9). Note that (9) depends negatively on how government values marginal consumption of top earners, on the elasticity of taxable income of top earners, and on the Pareto parameter. Saez (2001) explains the intuition behind this Pareto parameter result: "if the distribution is thin then raising the top rate for high income earners will raise little additional revenue" (Saez (2001), pp. 212).

Table 6
Top marginal income tax rates

		Elasticity of Taxable Income	
		0.46	0.76
		(1)	(2)
	Social Marginal Utility g= 0		
	Pareto Parameter		
(1)	1.4	60.8%	48.4%
	1.5	59.2%	46.7%
	2.0	52.1%	39.7%
	2.4	47.2%	35.1%
	Social Marginal Utility g= 0.1		
	Pareto Parameter		
(2)	1.4	58.3%	45.8%
	1.5	56.6%	44.1%
	2.0	49.5%	37.2%
	2.4	44.6%	32.8%

This table displays calculations of the marginal income tax rate for top earners (tax payers with quarterly incomes above $65,000 pesos) using formula (9). Optimal top MITRs are calculated for different values of the elasticity of taxable income, the Pareto parameter, and when government assigns different values to the marginal consumption of high income taxpayers.

As shown in Table 6, the Pareto parameter has a large influence on optimal tax rates. Estimations I made of this parameter are partly obtained from data on the United States as a consequence of the method I used to inflate top incomes in ENIGH. A more precise estimation of the parameter would be obtained from Mexican tax return statistics. As previously discussed, this information is not available for public use in Mexico. In light of the most likely values of the Pareto parameter (around 1.4 and 2), and of the elasticity values obtained from international estimates, marginal income tax rates for top earners should range between 45 and 60 percent. It is important to note that, as when optimal marginal tax schedules were calculated, these tax rates assume that

there are no other taxes that affect the choice between tax-favored and non-tax favored consumption and compensation.

However, there are two problems that affect the optimal MITRs calculated for the whole income schedule that do not affect the calculations of top marginal income tax rates. The first is the endogeneity problem: the income distribution is affected by taxation, thus, a might depend on the marginal tax rate, however, Saez (2001)[55] shows that a is independent of T' as long as $T' < 1$ and the ETI of top income taxpayers does not vary among them. Thus one can use the empirical value of a to get T' without having endogeneity problems. The other problem is the one concerning the informal sector: since the largest share of top incomes is part of the formal sector, if the process I undertook to inflate high incomes in ENIGH is valid, and by valid I mean that it gives an appropriate value of a, then informality of Mexican economy should not be a problem since it is mostly present in low paid work force.

Due to the reasons just explained, optimal MITR calculations for top earners using (9) are preferred to those made using the optimal schedule formula (8). Thus, if assumptions on the ETIs and the Pareto parameter are realistic, marginal income tax rates for Mexican top earners should be around 45 and 60 percent. This means: taxpayers with quarterly incomes above $65,000 pesos should pay between 45 and 60 percent in taxes on any additional income.

Saez (2001) performs calculations for the MITR for top earners in the United States, he gets higher rates: his preferred estimate range is between 60 and 70 percent.[56] This is mainly due to the fact that his elasticity estimates are smaller. Comparison of my results with Saez's results highlight how important is the thinness of the top tail of the income distribution and top earners' ETI for determining the optimal top MITR: the thinner the distribution and the bigger their response, the less additional revenue government can get from raising top tax rates to top earners.

[55] See Saez (2001) for details on the proof.
[56] Diamond and Saez (2011) set the optimal MITR for top earners in the United States at 73 percent; they get this rate using an ETI of 0.25. Fieldhouse (2013) argues that a 73 percent MITR implies that, combining state and federal taxes, the current top MITR in the Unites States is 27.5 percent below the optimal maximizing level.

5.5. *Some considerations on income taxation, growth and inequality*

I would like to close this Section commenting the implications that income taxes have on economic growth and inequality. MITRs derived in this work may appear challenging to most readers as a consequence of how high they can get. Thought this work, we have seen that high income tax rates are consistent with taxpayers' behavioral responses; however, there is a widespread concern that income taxes may discourage economic growth. Many authors[57] argue that this concern may be the result of supply side economics' postulate that income or capital taxes may discourage savings and investment, and this will in turn decrease growth. This argument applies mostly to top earners since generally they drive investment in the economy. In fact, this may have been a reason why many countries dropped income tax rates for top earners in the 1970's and the 1980's.

If increasing top income taxes has a negative effect on economic growth, then it would be a bad idea to increase them since long-term welfare for the whole population would be compromised. However, several authors have performed statistical analysis to find if income taxes, particularly income taxes for top earners, are negatively correlated with economic growth. What researchers have found is that there is no correlation between income taxation and economic growth. Piketty et al. (2011) perform regression analysis for the period of 1960 to 2010, they find no strong correlation between top income taxation and economic growth. Hungerford (2012) also performs empirical analysis to find correlation between taxes for top earners and growth; he finds no correlation between those variables, neither he finds correlation between top income taxes and growth drivers such as savings, investment and productivity. Gravelle and Marples (2011) do not find relation between top income taxes and growth, savings or labor participation. Romer and Romer (2012) find no significant impact of marginal income tax rates on investment or construction.

Piketty et al. (2011) argue that top income taxation not being correlated to growth is consistent with the hypothesis that a part of gains for top earners come at the expense of lower earners. They find evidence that top earners have considerable bargaining power in determining their incomes, a large part of their compensation is not due to their marginal product but to their

[57] See Diamond and Saez (2011), Piketty et al. (2011), Fieldhouse (2013), Gravelle and Marples (2011) or Hungerford (2012).

ability to bargain. They also find that higher income taxes limit top earners' bargain power. Thus high top taxes not only serve to increase revenue collection from top earners, but also to impede top earners augmenting their income at the expense of lower earners.

Moreover, statistical analysis finds correlation between top income tax rates and income inequality. Hungerford (2012) finds that when top taxes rates are reduced, top earners get a larger share of total income. Piketty et al. (2011) find similar results. Hungerford (2011) examines the relation between the Gini coefficient and top income tax rates, he finds evidence that increasing taxes for top earners can reduce the Gini coefficient. My numerical example of Section 5.3 points in the same direction.

In sum, research shows that, if taxpayers' behavioral responses and their income distribution allow it, high income tax rates, particularly for top earners, are desirable, since they increase revenue collection without affecting economic growth. Besides, high income tax rates for top earners serve as a tool to increase social welfare since they can potentially increase income equality in a society. High marginal income taxes for top earners have existed in industrialized countries in the past, after World War II, top marginal income taxes reached 90 percent in many countries, in the same period of time, industrialized countries experienced high growth rates. Thus, empirical evidence and historical experience gives strong arguments in favor of high marginal income tax rates for top earners.

6. CONCLUSION

Determining optimal marginal income tax rates using the models developed by Mirrless (1971) and by Saez (2001) is a fruitful exercise because it allows government to determine rates that maximize revenue given taxpayers' behavioral responses and their income distribution. However, applying this theoretical framework to Mexico has several limitations that impede calculating in a precise manner marginal income tax rates that would maximize revenue. Due to the lack of relevant data, rates obtained in this work should not be taken as definitive marginal income tax rates that should be applied to Mexican taxpayers. This work highlights how important is that Mexican tax authorities guarantee public access to data on individual tax

returns. Without the needed data, definition of income tax rates could be defective, and may even be harmful for the economy.

Despite limitations in the data, I obtain four main conclusions from the calculations performed in this work. 1) The optimal marginal income tax schedule that maximizes collected revenue follows a U-shape pattern, where tax rates should be higher for low and high earners, and lower for middle earners. 2) There are large positive effects on the well-being of poor individuals and on income equality from an income tax system that redistributes collected revenue. 3) The Pareto parameter of Mexican income distribution has an upper bound value of 2.4, and its value is most likely between 1.5 and 2. 4) Given international ETI estimates and realistic values of the Pareto parameter, marginal income tax rates for top earners in Mexico should be between 45 and 60 percent.

There are several extensions that can be performed to the analysis made in this work. One obvious extension would be deriving optimal MITRs using elasticity estimates and distribution of taxpayers' incomes obtained from Mexican individual tax return data, for the time being, available data does not allow performing this extension. When tax return data is available, obtaining values for the Pareto parameter of Mexican income distribution would be a necessary extension. Another extension could be using different income distributions to calculate optimal MITR schedules, I performed calculations using ENOE that I do not show in this work, I get a U-shape pattern but the level of tax rates changes, and this indicates that tax rates are very sensitive to the shape of the income distribution. A different extension could be analyzing in a more detailed manner the effects in the income distribution caused by redistributing collected revenue, the impacts of redistribution on top incomes is not analyzed in this work and it could provide interesting results. Another interesting extension could be using Piketty's et. al (2011) top income marginal taxation model that incorporates not only taxable income elasticities, but also tax avoidance and bargaining elasticities, however, this analysis can be carried out only until tax return data on Mexican top earners be open to public use.

REFERENCES

Atkinson, Anthony B. 1995. *Public Economics in Action: The Basic income/flat Tax Proposal*: Oxford University Press.

Atkinson, Anthony B. and Andrew Leigh. 2008. "Top Incomes in New Zealand 1921-2005: Understanding the Effects of Marginal Tax Rates, Migration Threat, and the Macroeconomy" *Review of Income and Wealth*, 54(2): 149-165.

Atkinson, Anthony B. and Thomas Piketty (Eds.). 2007**.** *Top Incomes Over the Twentieth Century: A Contrast between Continental European and English-Speaking Countries*: Oxford University Press.

Atkinson, Anthony B., Thomas Piketty, and Emmanuel Saez. 2011. "Top Incomes in the Long Run of History" *Journal of Economic Literature*, 49(1): 3-71.

Austan Goolsbee. 2000. "What Happens when You Tax the Rich? Evidence from Executive Compensation" *Journal of Political Economy*, 108(2): 352-378.

Auten, Gerald and Robert Carroll. 1999. "The Effect of Income Taxes on Household Income" *Review of Economics and Statistics*, 81(4): 681-693.

Auten, Gerald, Robert Carroll, and Geoffrey Gee. 2008. "The 2001 and 2003 Tax Rate Reductions: An Overview and Estimate of the Taxable Income Response" *National Tax Journal*, 61(3): 345-364.

Bakos, Péter, Péter Benczúr, and Dora Benedek. 2008. "The Elasticity of Taxable Income: Estimates and Flat Tax Predictions using the Hungarian Tax Changes in 2005" *EUI Working Papers* (RSCAS 2008/32).

Blomquist, Sören and Håkan Selin. 2010. "Hourly Wage Rate and Taxable Labor Income Responsiveness to Changes in Marginal Tax Rates" *Journal of Public Economics*, 94(11–12): 878-889.

Bloom, David E., David Canning, and Jaypee Sevilla. 2004. "The Effect of Health on Economic Growth: A Production Function Approach" *World Development*, 32(1): 1-13.

Brewer, Mike, Emmanuel Saez, and Andrew Shephard. 2010. "Means Testing and Tax Rates on Earnings" *Dimensions of tax Design: the Mirrlees Review*: 90-173.

Diamond, Peter and Emmanuel Saez. 2011. "The Case for a Progressive Tax: From Basic Research to Policy Recommendations" *Journal of Economic Perspectives*, 25(4): 165-190.

Díaz Mendoza, Manuel. 2004. "La Respuesta De Los Contribuyentes Ante Las Reformas Del IRPF, 1987-1994" *Tesina del Centro de Estudios Monetarios y Financieros*.

Feldstein, Martin. 1995. "The Effect of Marginal Tax Rates on Taxable Income: A Panel Study of the 1986 Tax Reform Act" *Journal of Political Economy*, 103(3): 551-572.

------. 1999. "Tax Avoidance and the Deadweight Loss of the Income Tax" *The Review of Economics and Statistics*, 81(4): 674-680.

Fieldhouse, Andrew. 2013. "A Review of the Economic Research on the Effects of Raising Ordinary Income Taxes" http://www.epi.org/files/2013/raising-income-taxes.pdf.

Flores, Ernesto. 1975. "Trayectoria Del Impuesto Sobre La Renta En México" *Revista de la Facultad de Derecho de México*, 99-100: 627-662.

Giertz, Seth H. 2007. "The Elasticity of Taxable Income Over the 1980s and 1990s" *National Tax Journal*, 60(4): 743-768.

Goolsbee, Austan, Robert E. Hall, and Lawrence F. Katz. 1999. "Evidence on the High-Income Laffer Curve from Six Decades of Tax Reform" *Brookings Papers on Economic Activity*, 1999(2): 1-64.

Gottfried, Peter and Hannes Schellhorn. 2004. "Empirical Evidence on the Effects of Marginal Tax Rates on Income--the German Case" http://www.eea-esem.com/papers/eea-esem/2004/2562/elasticity_of_taxable_income.pdf .

Gravelle, Jane and Donald Marples. 2011. "Tax Rates and Economic Growth" http://www.fas.org/sgp/crs/misc/R42111.pdf.

Gruber, Jon and Emmanuel Saez. 2002. "The Elasticity of Taxable Income: Evidence and Implications" *Journal of Public Economics*, 84: 1-32.

------. 2000. "The Elasticity of Taxable Income: Evidence and Implications" *NBER Working Paper No. 7512*.

Hansson, Åsa. 2007. "Taxpayers' Responsiveness to Tax Rate Changes and Implications for the Cost of Taxation in Sweden" *International Tax and Public Finance*, 14(5): 563-582.

Heckman, James J. 1993. "What has been Learned about Labor Supply in the Past Twenty Years?" *American Economic Review*, 83(2): 116-121.

Hungerford, Thomas. 2011. "Changes in the Distribution of Income among Tax Filers between 1996 and 2006: The Role of Labor Income, Capital Income, and Tax Policy" http://taxprof.typepad.com/files/crs-2.pdf.

------. 2012. "Taxes and the Economy: An Economic Analysis of the Top Tax Rates since 1945" http://graphics8.nytimes.com/news/business/0915taxesandeconomy.pdf.

Kleven, Henrik J. and Esben A. Schultz. 2011. "Estimating Taxable Income Responses using Danish Tax Reforms" *EPRU Working Paper Series 2011-02*.

Kleven, Henrik J. and Mazhar Waseem. 2011. "Tax Notches in Pakistan: Tax Evasion, Real Responses, and Income Shifting" http://sticerd.lse.ac.uk/seminarpapers/pes16112011.pdf.

Kopczuk, Wojciech. 2010. "Polish Business Flat Tax and its Effect on Reported Incomes: A Pareto Improving Tax Reform?" http://www.columbia.edu/~wk2110/bin/PolishFlatTax.pdf.

Krueger, Alan B. and Mikael Lindahl. 2001. "Education for Growth: Why and for Whom?" *Journal of Economic Literature*, 39(4): 1101-1136.

Levy, Santiago. 2008. *Good Intentions, Bad Outcomes: Social Policy, Informality, and Economic Growth in Mexico*: Brookings Institution Press.

Lindert, Peter H. 2004. *Growing Public: Volume 1, the Story: Social Spending and Economic Growth since the Eighteenth Century*. Vol. 1: Cambridge University Press.

Lindsey, Lawrence B. 1987. "Individual Taxpayer Response to Tax Cuts: 1982–1984: With Implications for the Revenue Maximizing Tax Rate" *Journal of Public Economics*, 33(2): 173-206.

MaCurdy, Thomas. 1992. "Work Disincentive Effects of Taxes: A Reexamination of some Evidence" *The American Economic Review*, 82(2, Papers and Proceedings of the Hundred and Fourth Annual Meeting of the American Economic Association): 243-249.

Mirrlees, James A. 1971. "An Exploration in the Theory of Optimum Income Taxation" *Review of Economic Studies*, 38(114): 175-208.

Pencavel, John H. 1986. "Labor Supply of Men: A Survey." In *Handbooks in Economics Series, no. 5*, ed. Orley Ashenfelter and Richard Layard, 3-102.

Piketty Thomas. 1999. "Les Hauts Revenus Face Aux Modifications Des Taux Marginaux Supérieurs De l'Impôt Sur Le Revenu En France, 1970-1996" *Économie & prévision*, 138-139: 25-60.

Piketty, Thomas and Emmanuel Saez. 2007. "How Progressive is the U.S. Federal Tax System? A Historical and International Perspective" *Journal of Economic Perspectives*, 21(1): 3-24.

------. 2012. "Optimal Labor Income Taxation" *NBER Working paper No. 18521*.

Piketty, Thomas: S., Emmanuel and Stefanie Stantcheva. "Optimal Labor Income Taxation: A Tale of Three Elasticities" *NBER Working paper No. 17616*.

Pirttila, Jukka and Hakan Selin. 2006. "How Successful is the Dual Income Tax? Evidence from the Finnish Tax Reform of 1993" *CESifo Working Paper Series No. 1875*.

Romer, Christina D. and David H. Romer. 2012. "The Incentive Effects of Marginal Tax Rates: Evidence from the Interwar Era" *NBER Working Paper No. 17860*.

Saez, Emmanuel. 2001. "Using Elasticities to Derive Optimal Income Tax Rates" *The Review of Economic Studies*, 68(1): 205-229.

Saez, Emmanuel, Joel Slemrod, and Seth H. Giertz. 2009. "The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review" *NBER Working paper 15012*.

------. 2012. "The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review" *Journal of Economic Literature*, 50(1): 3-50.

Saez, Emmanuel and Michael R. Veall. 2005. "The Evolution of High Incomes in Northern America: Lessons from Canadian Evidence" *The American Economic Review*, 95(3): 831-849.

Salanié, Bernard. 2003. *The Economics of Taxation*: The MIT press.

Sierra, Carlos J. 1959. "Historia De La Secretaría De Hacienda y Crédito Público II" *Revista de Administración Pública*, 11: 55-67.

Sillamaa, Mary-Anne and Michael R. Veall. 2001. "The Effect of Marginal Tax Rates on Taxable Income: A Panel Study of the 1988 Tax Flattening in Canada" *Journal of Public Economics*, 80(3): 341-356.

Slemrod, Joel. 1994. "Fixing the Leak in Okun's Bucket Optimal Tax Progressivity when Avoidance can be Controlled" *Journal of Public Economics*, 55(1): 41-51.

------. 1996. "High-Income Families and the Tax Changes of the 1980s: The Anatomy of Behavioral Response." In *Empirical Foundations of Household Taxation*, ed. Martin Feldstein and James Poterba: University of Chicago Press.

------. 1998. "Methodological Issues in Measuring and Interpreting Taxable Income Elasticities" *National Tax Journal*, 51(4): 773-788.

Tuomala, Matti. 1990. "Optimal Income Tax and Redistribution" *Oxford University Press*.